

Article

# Marine invasions enter the genomic era: three lessons from the past, and the way forward

Frédérique VIARD<sup>a,\*</sup>, Patrice DAVID<sup>b</sup>, and John A. DARLING<sup>c</sup>

<sup>a</sup>Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7144, Lab. Adaptation Et Diversité En Milieu Marin, Team Div&Co, Station Biologique De Roscoff, Roscoff 29682, France, <sup>b</sup>CEFE UMR 5175, CNRS-Université De Montpellier-UM III-EPHE, 1919 Route De Mende, Montpellier Cedex 34293, France, and <sup>c</sup>National Exposure Research Laboratory, United States Environmental Protection Agency, Research Triangle Park, NC 27711, USA

\*Address correspondence to Frédérique Viard. E-mail: viard@sb-roscoff.fr.

Received on 27 January 2016; accepted on 15 April 2016

## Abstract

The expanding scale and increasing rate of marine biological invasions have been documented since the early 20th century. Besides their global ecological and economic impacts, non-indigenous species (NIS) also have attracted much attention as opportunities to explore important eco-evolutionary processes such as rapid adaptation, long-distance dispersal and range expansion, and secondary contacts between divergent evolutionary lineages. In this context, genetic tools have been extensively used in the past 20 years. Three important issues appear to have emerged from such studies. First, the study of NIS has revealed unexpected cryptic diversity in what had previously been assumed homogeneous entities. Second, there has been surprisingly little evidence of strong founder events accompanying marine introductions, a pattern possibly driven by large propagule loads. Third, the evolutionary processes leading to successful invasion have been difficult to ascertain due to faint genetic signals. Here we explore the potential of novel tools associated with high-throughput sequencing (HTS) to address these still pressing issues. Dramatic increase in the number of loci accessible via HTS has the potential to radically increase the power of analyses aimed at species delineation, exploring the population genomic consequences of range expansions, and examining evolutionary processes such as admixture, introgression, and adaptation. Nevertheless, the value of this new wealth of genomic data will ultimately depend on the ability to couple it with expanded “traditional” efforts, including exhaustive sampling of marine populations over large geographic scales, integrated taxonomic analyses, and population level exploration of quantitative trait differentiation through common-garden and other laboratory experiments.

**Key words:** biological invasions, cryptic species, cryptogenic species, eco-evolutionary processes, genomics, metabarcoding.

## Introduction

Invasive species are widely recognized for their negative ecological and economic impacts, and have been implicated as one of the major drivers of global biodiversity decline (Brook et al. 2008). In part this reflects direct negative impacts of introduced species on native populations through competitive, predatory, or other ecological interactions (Simberloff 2005). More subtle but equally destructive is the trend toward biotic homogenization at regional and even global

scales, leading many observers to recognize earth’s entry into an unprecedented new era in biogeographic history (Mooney and Cleland 2001). At the same time, biological invasions frequently provide opportunities to glean insights into important biological processes associated with global change (Moran and Alexander 2014). The ability of populations to cope with rapidly changing environmental conditions is a central and timely question in evolutionary ecology, particularly given observations of ongoing ecological and

evolutionary changes in response to anthropogenic stressors at multiple spatial and temporal scales (Parmesan 2006; Poloczanska et al. 2013). The deliberate or accidental translocation of populations outside their native range by human activities continues to occur across extended latitudinal ranges, oceans, and ecosystems (Molnar et al. 2008). The ability of certain populations to establish and spread in biotic and abiotic environments in which they have no evolutionary history—the so-called “paradox of invasion” (Sax and Brown 2000)—suggests that study of these systems may deepen our understanding of the biotic interactions, demographic patterns, and eco-evolutionary processes leading to success in the face of rapidly changing environmental conditions.

A review of literature published since 1990 suggests that population genetics and related scientific fields (e.g., phylogeography, parentage analyses, molecular barcoding, quantitative genetics) have become popular approaches for understanding patterns and mechanisms associated with establishment of non-native species (Viard and Comtet 2016). The questions addressed by means of these genetic tools are diverse and apply to all stages of the invasion process. Marine non-native species are no exception, with important progress made during the past 25 years, as highlighted in a number of recent reviews (Roman and Darling 2007; Geller et al. 2010; Rius et al. (2015); Viard and Comtet 2016); however, genetic studies of marine invaders are nevertheless considerably less common than those of their terrestrial and freshwater counterparts. A literature survey by Rius et al. (2015) of all studies published to 2013 reported fewer than 60 taxa that have been introduced to Europe for which population genetics analyses *sensu lato* were available, representing only a small fraction of the 1369 introduced species reported in European marine waters (Katsanevakis et al. 2013).

Although genetic studies of marine non-native taxa are relatively uncommon, the taxa that have been investigated are diverse enough (e.g., molluscs, tunicates, algae, etc.) to highlight several consistent findings. In particular, 3 observations have been consistently reported in marine invasion genetic studies:

1. There is a surprisingly high number of cryptic and/or cryptogenic species among marine taxa. Genetic studies, particularly those carried out to determine invader origins, frequently uncover previously unrecognized diversity. These findings have implications for our ability to reconstruct historical species assemblages and understand ecological processes acting at community levels.
2. Founder events in introduced populations are far from being the rule when considering non-native marine taxa. This result has important consequences for understanding the processes of establishment and identifying the drivers by which a species may become invasive. The relationship between genetic diversity and the success of introduced marine populations thus remains uncertain.
3. It remains unclear to what degree adaptation drives the success of introduced populations. To date, only few studies have directly addressed the role that adaptive processes may play in the course of marine invasions. Exploration of these processes is critical not only for understanding the evolutionary context of marine invasions, but also to evaluate the extent to which marine invasions may serve as case studies for rapid evolutionary response to environmental change.

The vast majority of genetic studies addressing marine invasions have to this point been based on what we will refer to here as “traditional” tools and techniques (e.g., Sanger sequencing, phylogenetic

reconstructions using standard mitochondrial loci, population genetic inference based on a small number of microsatellite loci, etc.). In this article, we wish to explore the potential for rapidly emerging new approaches based primarily on high-throughput sequencing (HTS) to address the issues outlined above. Such approaches are increasingly being proposed as means to overcome limitations imposed by traditional methods. HTS-based population genomics analyses offer the opportunity to query simultaneously thousands of markers spread over the genome, and thus potentially address a larger spectrum of ecological and evolutionary processes previously inaccessible (Ellegren 2014). They also provide tools capable of characterizing entire biological communities with unprecedented sampling throughput (Cristescu 2014). As illustrated elsewhere in this special issue, these methods are already yielding new insights into the evolutionary ecology of marine species. Here we explore specifically their application to marine invasions, with particular emphasis on the potential for these new tools to overcome existing methodological limitations to addressing the issues highlighted above.

### Understanding Patterns of Marine Invasions: Resolving Cryptic and Cryptogenic Taxa

It is very likely that we have vastly underestimated the degree to which anthropogenic introductions have reshaped marine biogeography (Geller et al. 2010). The large number of cryptogenic species in marine systems is only 1 indication that translocation by human activities has likely shuffled marine biodiversity for centuries without notice. In addition, it is certainly the case that a substantial proportion of introduced species remain undetected (Carlton 2009). This may be because those species are secretive, establishing at low densities in poorly studied habitats or belonging to taxa that present challenges to traditional survey methods, or because they are observed but mistaken for native species. Furthermore, the high frequency with which known invasive marine species are revealed to be closely related complexes of evolutionary lineages (many likely independent cryptic species themselves) provides additional indication that we have yet to fully appreciate the role of invasions in altering patterns of marine biodiversity. In all such cases, the application of genetic and genomic tools may considerably enhance our ability to resolve contemporary biogeographic patterns, and to infer historical changes in such patterns.

Perhaps the most obvious application of these tools lies in delineating current distributions of invasive marine species. Unfortunately, traditional approaches to species distribution monitoring exhibit limitations that may render them insufficiently robust to support such efforts (Pfrender et al. 2010). Classical biodiversity inventories have relied on expert identifications of species based on morphological taxonomy. Such methods continue to be invaluable, but they are labor intensive and therefore often slow and expensive, and they rely on a dwindling institutional capacity for expert taxonomic identification. In addition, they are limited by an inability to identify individuals for which diagnostic morphological characters may be absent, such as subadults or members of species lacking useful identification keys (Besansky et al. 2003). These challenges are even more pronounced for certain important and speciose groups that resist morphological description, such as eukaryotic meiofauna (Creer et al. 2010).

For the past decade, DNA barcoding has been offered as a partial solution to the problems facing traditional biodiversity assessments (Hebert et al. 2003). Despite acknowledged theoretical and practical

limitations, DNA barcoding has been widely adopted to improve species-level identifications of individual organisms and to highlight possible cases of previously undescribed diversity, in both research and environmental management contexts. More recently, the potential value of this approach has been expanded dramatically through marriage with HTS (Cristescu 2014). Metabarcoding—the simultaneous barcoding of large numbers of individuals collected in a single complex environmental sample—offers the possibility of biodiversity assessments with processing throughput that far outstrips the ability of traditional approaches based on morphological identification (Taberlet et al. 2012; Comtet et al. 2015).

These approaches can support efforts to better understand distributions of introduced species in 2 different ways. First, they offer enhanced efficiency of detection over traditional methods, particularly for rare and hard-to-detect species. This, combined with the expected cost effectiveness of large scale metabarcoding studies compared with standard monitoring efforts, could lead to dramatic increases in the number of observations and thus far more thorough descriptions of species distributions. Second, the ability to discriminate between cryptic species or evolutionary lineages provides taxonomic resolution typically unavailable to traditional approaches, potentially resulting in more accurate and more detailed distributional data. A number of recent studies have begun to demonstrate these benefits at multiple spatial scales. For instance, Pochon et al. (2013) recently determined that metabarcoding based on the Cytochrome c oxidase I (COI) locus could detect larvae of the widely invasive seastar *Asteria amurensis* in environmental samples with sensitivity levels comparable to quantitative Polymerase Chain Reaction (qPCR) assays, and Zhan et al. (2013) demonstrated high-sensitivity detection of rare species, including known invasive species, artificially introduced into complex plankton communities. In 1 notable recent study, Zaiko et al. (2015a) showed that metabarcoding efforts could detect invasive species in the Baltic Sea far more effectively than could traditional monitoring approaches. The same authors similarly revealed the potential of metabarcoding to examine biodiversity being actively translocated via known vectors of anthropogenic introduction, such as ballast water (Zaiko et al. 2015b). Further expanding the possible utility of these methods are a number of studies suggesting that accurate surveys for marine taxa can be conducted using environmental DNA (eDNA) in marine systems (Thomsen et al. 2012; Kelly et al. 2014). Overall, there is reason to believe that such approaches may be particularly valuable in marine and other aquatic environments, where traditional survey methods can be especially challenging and the capture of eDNA may be especially likely.

These results all illustrate the capacity of metabarcoding approaches to detect non-native species even for rare taxa and for difficult to identify meroplankton species (Thomsen et al. 2012; Lindeque et al. 2013), while simultaneously gathering information on overall community diversity, including for taxonomic groups previously shielded from inquiry. Such efforts stand to provide critically important assistance in resolving the invasive status of many taxa. Of particular note is the capacity of metabarcoding to facilitate so-called “passive surveillance” (Simmons et al. 2016). Although genetic tools are often applied to understand species distributions via targeted monitoring (most frequently through species-specific Polymerase Chain Reaction (PCR) or other probe-based methodologies), metabarcoding provides opportunity to gather information on taxa not identified a priori as surveillance targets. Thus, not only would the broad application of metabarcoding efforts provide more accurate estimates of the current distributions of known invasives, it

could potentially uncover and provide detailed distribution information on previously unrecognized invasive species. All of this is crucial, as uncertainty in known distribution accounts for much confusion surrounding many species of indeterminate provenance, particularly when questions arise regarding putative native ranges or disjunctions in observed global distributions (Carlton 2009). For instance, of 10 criteria adopted by Chapman and Carlton (1991) to determine introduced status of a species, 4 of them require detailed knowledge of that species’ distribution. False or incomplete knowledge of distribution could seriously impair assessments of species status and even result in mistaken identification of species as native or invasive. More thorough understanding of species distributions driven by metabarcoding studies thus has the potential to reduce substantially the high frequency of cryptogenic and pseudoinigenous species (species for which native origins are either unknown or mistaken) that persists among marine taxa (Carlton 2009).

Despite the promise held by metabarcoding approaches, there remain a number of pressing questions regarding their ultimate utility for surveillance of marine invasions. The primary challenge lies in the difficulty of interpreting sequences that appear at very low frequencies in metabarcoding datasets. Singleton sequences are especially problematic: If a taxon is represented by only a single sequence, is that a true reflection of that species’ rarity in the environmental sample, or is it simply an error that should be discarded when processing the raw sequence data? One study of microbial populations found that diversity of rare species could be grossly overestimated unless very stringent filtering protocols were adopted to remove likely sequencing errors (Kunin et al. 2010). However, Zhan et al. (2014) also recently demonstrated that greater stringency in the removal of presumed artifacts eliminated sequences representing rare species present in the sample, with the effect being most pronounced for those species with lowest relative biomass. Future solutions to this problem will likely involve strategies that employ internal and external references to establish baselines for the effects of artifacts, the development of advanced algorithms for error detection in processing HTS datasets, and the use of multiple barcode loci to provide independent estimates of diversity from a single sample.

Estimates of abundance from HTS data also remain problematic. Recent studies suggest that at least rough abundance estimates can already be generated with metabarcoding data (Kelly et al. 2014; Elbrecht and Leese 2015). Despite these advances, it appears that metabarcoding approaches may in fact have intrinsic limitations in this regard. At the very least, differences across species in body size, cellular density, and mitochondrial copy number, along with variability in PCR amplification efficiency, will render it very difficult metabarcoding studies to achieve the abundance standard—absolute abundance based on individual counts—adopted in many traditional biodiversity monitoring contexts. In addition to further advances in abundance estimation, the utility of metabarcoding may therefore also depend on development of biodiversity indices based on presence/absence data alone (e.g., Aylagas et al. 2014) or the implementation of rigorously designed sampling strategies that allow indirect inference of target population densities underlying patterns of positive detections (e.g., Jerde et al. 2011).

A further challenge to the wide application of metabarcoding is the sufficiency of reference libraries. Contrary to the ideal of a single barcode for all species, numerous studies have now indicated that different loci are necessary for different taxa, and suggest that multiple loci will likely be necessary to interpret biodiversity in complex communities (Cristescu 2014). Furthermore, adequacy of data can vary widely across taxa and across study locations, and in many

cases may be insufficient to support direct assignment at the species level (Trebitz et al. 2015). These findings imply that full assignment of metabarcode sequences to species level will be impossible with currently available databases.

Nevertheless, there are several reasons to believe that metabarcoding studies may yet prove valuable for understanding distributions of non-native diversity. For one thing, useful information can be gathered even when only a fraction of barcode sequences can be assigned to known species. For instance, a recent assessment of reference libraries for the Great Lakes suggests that metabarcoding efforts will have a very high likelihood of accurately characterizing fish assemblages in that region, including both native and non-native species, even if a substantial proportion of non-fish barcodes yield no species-level taxonomic information (Trebitz et al. 2015). In addition, even when species-level matches are unavailable barcodes can correctly be assigned to higher level taxonomic groups (genus or family; Wilson et al. 2011), potentially facilitating assignment of non-native status even if species-level identification is unavailable. It is also clear that in some cases, sufficient human and financial resources exist to support rapid accumulation of reference sequences for high profile taxa or regions. Indeed, for some well-studied systems reference libraries may already be sufficiently robust to support active metabarcoding surveillance (Raupach et al. 2015).

It is worth noting that in the face of rapid ongoing improvements in metabarcoding technologies, in particular in relation to primer design (e.g., for targeting specific taxonomic groups) and bioinformatic pipelines, it is crucial to establish rigorous DNA sample archiving policies. Proper archiving of DNA collected from environmental samples will enable not only long-term longitudinal surveys but also possibly *a posteriori* identification of non-indigenous species (NIS) first undetected because of technical limitations. Although archived HTS sequence databases allow re-analysis based on new bioinformatics workflows, archived DNA could be explored using different loci capable of discriminating taxa unrecognized in initial studies (or newly identified via independent integrated taxonomic efforts), or could even be subjected to analyses based on novel technologies that generate more robust data than currently available. DNA samples are typically archived by individual researchers, as DNA is relatively amenable to inexpensive long-term storage. However, the utility of such collections is limited outside of the research groups that generated them, and at this time we are not aware of any broadly available, well curated archives of DNA generated from monitoring surveys.

As noted above, 1 persistent challenge in accurately assessing the distribution of introduced marine species is the problem of cryptic species. Marine systems are replete with cryptic species complexes (Appeltans et al. 2012), and examples of molecular studies uncovering such complexes are far too numerous to discuss in any detail here (Geller et al. 2010; Riesgo et al. 2015). The recognition of cryptic species is especially pervasive—and problematic—in the case of marine invasions, where uncertainty in the identities of introduced taxa can result in misguided management strategies and can obscure true introduction pathways and invasion pressures (McGlashan et al. 2008; Knapp et al. 2015). Unfortunately, the identification of cryptic lineages via phylogeographic or population genetic studies is typically made on the basis of molecular results from 1 or very few genetic loci (often standard barcoding loci such as COI and 18S), and should thus most prudently be taken as taxonomic hypotheses rather than definitive statements on species identity. The result is a literature scattered with untested hypothetical records of undescribed marine biodiversity.

Ideally, such hypotheses would be addressed through rigorous integrative taxonomic efforts that employ both DNA-based and traditional systematic approaches to delineate true species boundaries and inform systematic revisions (Pante et al. 2015a). For example, after the identification of 4 cryptic lineages within the invasive vase tunicate *Ciona intestinalis*, a taxonomic revision based on joint molecular and morphological analyses resulted in acceptance of 2 names for 2 of these lineages, namely *C. intestinalis* (formerly recognized as *C. intestinalis* type A) and *C. robusta* (formerly known as *C. intestinalis* type B; Brunetti et al. 2015). Unfortunately, this is a rather exceptional case among invasive marine species, reflecting unusually high investment in a species complex that has been adopted as a model system in various fields. For the majority of species without such resources, novel genomic methods driven by HTS may help to delineate species even in the absence of dedicated integrative efforts. The advent of HTS has already facilitated species delimitation studies by providing ready access to vast numbers of informative genomic loci even in non-model systems. Such recently developed tools as restriction site associated DNA (RAD) tags are already being used to test hypotheses of interspecific relationships and to clarify boundaries between cryptic lineages (Pante et al. 2015b). By allowing analysis of data from large numbers of unlinked genomic loci, these HTS-driven methods stand to overcome limitations of single-marker approaches to species delimitation, such as incongruence between gene and species trees caused by incomplete lineage sorting or other evolutionary phenomena. These options bring considerably greater power to DNA-based inferences of systematic relationships, and thus promise novel solutions in the integrative taxonomic toolkit. Future applications of these tools could aid considerably in clarifying the diversity of introduced marine lineages, thus providing greater insight into the degree to which invasions have reshaped marine communities.

## How Do Invasions Affect Genetic Diversity in Marine Organisms?

The invasion genetics literature published at the end of the 20th century revealed great hope in the use of polymorphic markers (e.g., microsatellites, Amplified Fragment Length Polymorphisms (AFLPs), Random Amplified Polymorphic DNAs (RAPDs)) to examine the genetic diversity of introduced populations and compare them to putative sources in the native range, in order to examine the importance of genetic drift and the consequences of founder events during the introduction and colonization stages (Sakai et al. 2001). Introduced populations are reasonably assumed to be seeded by only a small number of organisms, and a robust theoretical framework predicts that a reduction in population size leaves distinct genetic footprints such as reduction in overall genetic diversity, number of alleles or average heterozygosity, with variable magnitude depending on the size of the bottleneck and subsequent population growth rate (Nei et al. 1975). In non-native species, these founder effects are expected to influence the success of the invasion and possibly to explain the widely observed phenomenon of a post-establishment lag phase (Roman and Darling 2007; Dlugosch and Parker 2008; Bock et al. 2015). However, a number of studies have suggested that various mechanisms can lead to violations of that expectation, including multiple and repeated introductions as well as individual introductions with extremely high propagule loads. When a large numbers of individuals are introduced, either from a single diverse source or from a collection of multiple sources, the likelihood increases of introducing genotypes that exhibit high fitness under the ecological conditions found in the new range (i.e., “pre-



adapted” genotype), or alternatively of providing sufficient substrate on which selection may act (i.e., selection on standing genetic variation (Barrett and Schluter 2008); see Section “How Common—and How Important—Is Adaptation during the Invasion Process?”). In addition, with repeated introductions from multiple sources, new genetic combinations (“evolutionary novelties”) could arise from admixture resulting from intraspecific crosses between individuals with different genetic backgrounds (Rius et al. 2015).

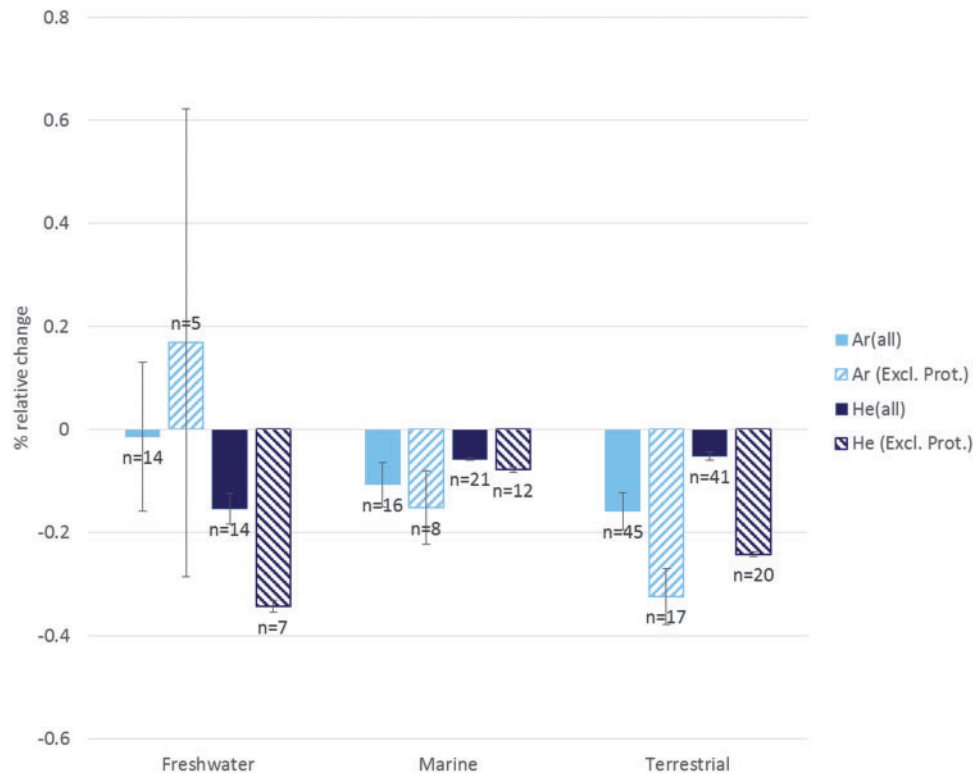
Examining the relationship between propagule pressure, the demography of introduction events, and the genetic consequences of founder events has thus been a central feature of invasion genetics. Propagule pressure has been proposed as a main factor contributing to the success of invasive species notably because an increase in number of founders is expected to decrease demographic stochasticity, reduce negative Allee effects, and increase the likelihood of introducing pre-adapted genotypes (Lockwood et al. 2005; Simberloff 2009). This hypothesis is of particular importance in marine systems, given that many marine species display life-history traits that may favor transport and release of large numbers of propagules. For example, the vast majority of the marine invertebrates (>70%; Mileikovsky 1971) are characterized by a benthopelagic cycle, with a free floating dispersive stage released in large quantities (typically, more than a few thousands of larvae per female). This is exemplified by the mollusc *Crepidula fornicata*, native to North America and invasive in Europe; at each reproduction event, the female releases on average 10,000 to 20,000 larvae (Broquet et al. 2015). Repeated introductions in time, from a single or multiple sources, is also an important facet of propagule pressure in marine systems, as revealed by genetic analyses (Geller et al. 2010; Rius et al. 2015). *C. fornicata* is also well-illustrating this feature: this mollusc was first introduced at the end of the 19th century, probably with spats of the American oyster *Crassostrea virginica*, from NE America. Then, massive accidental introductions were documented in the 1970s with the intentional introduction of the Pacific oyster *Crassostrea gigas* from Asia and NW America (Riquet et al. 2013 and references herein). Furthermore, the small size of these invertebrates’ larvae (typically in the range of 50–500  $\mu\text{m}$ ) makes them almost undetectable and thus favors their accidental transport. Some vectors such as ballast water are particularly efficient in delivering a large quantity of propagules (>1000) and very high number of taxa (>100) in a single inoculum (Gollasch 2008). Further contributing to invasion success may be the availability of life history stages that can lie dormant for months to years, and remain able to germinate when returned to favorable conditions (e.g., Briski et al. 2011). Altogether, the extraordinary fecundity, complex life-cycle, and large population sizes of many marine species make them excellent candidates for the delivery of high propagule pressures (Roman and Darling 2007).

The high propagule pressure suspected for many marine organisms should lead to a large number of founders and consequently to substantial genetic diversity in the introduced populations of marine species. In agreement with this prediction, in their review of ca. 80 marine invasion genetic studies, Rius et al. (2015) showed that introduced populations in Europe most often display genetic diversity similar to populations from the native range. In other words, marine invaders appear to be genetically quite diverse. Little to moderate loss of genetic diversity has been also documented in many terrestrial and freshwater environments (Dlugosch and Parker 2008), but marine species show even less pronounced relative changes in genetic diversity between native and introduced populations, when excluding enzymatic markers (Figure 1). Such genetically highly diverse

invasive populations support the hypothesis that adaptation from standing genetic variation may be an important process in marine invasions (see Section “How Common—and How Important—Is Adaptation during the Invasion Process?”).

Continued uncertainty regarding the importance of genetic diversity in marine invasions is rooted in part in limitations of traditional genetic tools for disentangling the complex relationships between diversity, demography, and colonization success. There are a number of ways in which population genomic approaches built on HTS technology might advance our understanding of these relationships. One important opportunity lies in unravelling the mystery of the sources of marine introductions. Without accurate reconstruction of invasion history, it is impossible to reasonably assess changes in genetic diversity associated with the introduction process. Unfortunately, the use of population genetics to study historical patterns of invasion is often limited by the very low genetic structure characterizing many marine organisms (Geller et al. 2010; Gagnaire et al. 2015). Indices used to measure genetic structure between populations, like  $F_{st}$ -based estimates, frequently display very low values in marine organisms (Selkoe and Toonen 2011). For instance, using data from Weersing and Toonen (2009),  $F_{st}$  values (averaged across species) at microsatellites loci range from 0.0016 in Echinodermata to 0.18 in Cnidaria. The few exceptions in the literature documenting very high  $F_{st}$  are likely explained by cryptic species and problems of species delineation (Pante et al. 2015c). This low genetic structure is likely often due to the combined effect of large effective population size and large migration rate (see Editorial and Box 1 in Gagnaire et al. 2015 for detailed explanations) common to marine invaders, as shown with a few examples provided in Table 1.

HTS-based genotyping approaches (for instance, Genotyping-By-Sequencing [GBS] techniques) promise to help overcome these limitations by delivering hundreds to thousands of markers at once, and thus allow for a higher statistical power in detecting small genetic variations, particularly when taking advantage of rare variants (O’Connor et al. 2014) or by selecting subsets of informative loci. For instance, Benestan et al. (2015) used >10,000 Single Nucleotide Polymorphisms (SNPs) identified from RAD sequencing (a GBS technique; Davey and Blaxter 2011) to examine fine-scale genetic structure of the American lobster, a species for which previous genetic studies supported the view of panmixia over a large geographic range. These authors showed that the average assignment rates to lobster sampling sites increased from 60.2% to 80.8% when selecting the 500 and 3000 most differentiated SNPs, respectively, compared with an assignment success rate of only 8.9% with all loci. The efficiency of HTS-based studies to examine the genetic structure of marine NIS was shown in 1 recent study of the globally invasive marine crab *Carcinus maenas* (Tepolt and Palumbi 2015). Those authors illustrated how a panel of >10,000 SNPs could reveal significant differentiation between invasive populations that were not distinguished previously based on mtDNA or microsatellites. Although such approaches offer substantial advantages over traditional methods for reconstructing invasion history, significant hurdles still remain. For instance, studies based on very large numbers of markers, though they provide greater power to resolve genetic divergence, may also exhibit higher sensitivity to sampling issues and linkage disequilibrium (Waples 2015). In addition, HTS-based genotyping of neutral markers is unlikely to fully overcome the effect of large effective population size (Gagnaire et al. 2015). Finally, for identifying routes and sources, no genetic approach can avoid the requirement for adequate sampling in both the introduced and invasive range; this includes sampling as much as possible all putative



**Figure 1.** Relative change of allelic richness (Ar) and gene diversity (He) in introduced populations as compared to native populations, across different environments. Data (Ar and He values) from Dlugosch and Parker (2008) were used with similar data retrieved from relevant studies reviewed by Rius et al. (2015). The 2 dataset were merged. Relative changes of Ar and He were computed as explained in Dlugosch and Parker (2008) for each species and values averaged across species sorted in 3 categories (Freshwater, Marine, and Terrestrial). Mean values and standard errors are provided. The data used are all from proteins (enzymatic markers) and nuclear markers (AFLPs, RAPDs, SSRs, nuclear genes), thus excluding mitochondrial data. Plain bars correspond to all studies whatever the markers used and hatched bars stand for all studies excluding those with enzymatic markers. The number of species ( $n$ ) examined is indicated for each subset.

**Table 1.** Genome scans of marine NIS that include populations from the introduced and native ranges

Species	Number and type of markers	$n_{pop}$	$i_{pop}$	$F_{st}$	Reference
<i>Crepidula fornicata</i>	327 AFLP + 17 microsatellites	7	15	AFLPs NI : 0.029 NN : 0.037 II : 0.014 Microsatellites NI : 0.020 NN : 0.031 II : 0.004	Riquet et al. 2013
<i>Crassostrea gigas</i>	240 AFLP + 8 microsatellites + 30 SNP	1	15	NI (north) : 0.020 NI (south) : -0.002 NN : — I(south)I(north) : 0.022 I(south)I(south) : 0.001 I(north)I(north) : 0.013	Rohfritsch et al. 2013
<i>Carcinus menas</i>	10,809 SNPs from transcriptome sequencing	2	5	NI : 0.104 NN : 0.049 II : 0.033	Tepolt and Palumbi 2015

The number and type of marker used, as well as the number of populations from the native range ( $n_{pop}$ ) and introduced range ( $i_{pop}$ ) studied are indicated ;  $F_{st}$  indicates the average pairwise  $F_{st}$  between 2 native (nn), 2 introduced (ii), or 1 native and 1 introduced (ni) populations. In the *Crassostrea* case, introduced populations have been split into northern and southern populations (cf. text).

sources and develop a sampling strategy that takes into account the genetic structure in the native range (Geller et al. 2010).

In this context, the nature of the loci (i.e., neutral vs. linked to loci under selection) is a major factor to consider. When analyzing

thousands of markers spread over the genome, some of them may be linked to selected regions of the genome and behave non-neutrally, e.g., show an exceptionally high level of spatial differentiation. They are often called “selected” loci although in practice they are not

### Box 1—Patterns of adaptive evolution in introduced populations.


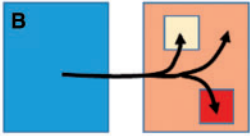
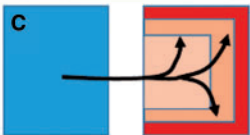

The simplest pattern (Figure 2A) involves a contrast between the source and recipient environments; adaptive differences will arise at this scale as a result of adaptation to general novel features of the new environment. However, the often considerable geographic extent of marine NIS makes it very likely that conditions will vary locally across both native and introduced ranges. In that case, adaptive evolution will take the form of ongoing adaptation to local environmental gradients (Figure 2B) and will tend to increase differentiation of selected traits or of loci under selection within the introduced range (potentially paralleling the structure of the native range). Other mechanisms of evolution during invasion do not rely on specific spatial structure of the environment. A contrast may develop between populations at the front of the expanding NIS population and more anciently invaded core populations (Figure 2C) if the foundation of new populations selectively filters phenotypes that favor colonization in open habitats; such phenotypes might later be replaced by others that are fitter in stable populations at carrying capacity, presumably evolving back toward the traits of native, non-expanding populations, as the invasion front progresses and front populations become core populations. This transient “evolution for invasiveness” is illustrated by cane toads in Australia, where higher leg lengths (providing higher dispersal) are found in front than in core populations (Phillips et al. 2006). However, front populations are also subject to other, non-adaptive transient effects. Indeed deleterious alleles might occasionally rise in frequency by chance as populations go through successive local founder events, creating a transient “expansion load” that is later compensated by immigration and homogenization within the general gene pool of the invaded range (Peischl et al. 2013; Peischl and Excoffier 2015). Finally, NIS may encounter related evolutionary lineages (at either the intra- or interspecific level) and start to exchange genes with them (Figure 2D). Each of the 4 situations (A, B, C, and D) has different temporal dynamics. In addition, they are expected to leave a distinct spatial signature in the form of a spatial contrast in the breeding values for some quantitative trait under selection or in allele frequencies at a selected locus.

shown or even assumed to be under direct selection. Using such outliers might be a promising approach as compared to simply increasing the number of neutrally behaving loci. The benefits of this approach have been pointed out by several authors in a biological conservation framework; Funk et al. (2012) advocated for a combined use of neutral and outlier markers for defining conservation units, in particular those that display different adaptive potential; Nielsen et al. (2012) showed that outliers may be particularly useful for assigning population of origins of fishes; and Gagnaire et al. (2015) proposed to use them to obtain genetically based inferences of dispersal in species with large effective population size. As noted above, a common theme in marine invasion genetics is the often weak traces of colonization history and low differentiation of invasive populations. Using markers influenced by selection could be particularly relevant in this situation, as the large effective

population size of marine species should strongly increase the efficiency of selection and the likelihood to detect outliers (Allendorf et al. 2010). For instance, native populations of the mollusc *C. fornicata* are so weakly genetically structured that inferring the source populations of its introduction in Europe was not possible using neutral markers, including with a large set of AFLPs and microsatellites (Table 1; Riquet et al. 2013). However, based on allelic frequencies obtained at 8 outliers identified with a genome-scan (see Table 1 and details in Section “How Common—and How Important—Is Adaptation during the Invasion Process?”), the southern part of the native range could confidently be excluded as a source of the European populations.

Population genomics also provides novel opportunities to explore directly the complex relationships between sequential founder events, genetic diversity, and the ecological success of colonizing populations. GBS combined with other HTS-based tools (e.g., RNA-seq) allow the joint analyses of changes in gene expression and allele frequency, at both neutral and non-neutral genes, at very early stages of colonization. Two facets of colonization dynamics are of particular importance at these early stages, even prior to longer-term adaptive responses (Section “How Common—and How Important—Is Adaptation during the Invasion Process?”): allelic changes in relation to stochastic processes, like those arising with gene surfing at expansion fronts; and processes associated with admixture. The former may be an important but underappreciated aspect of the invasion process. Recent empirical and theoretical work suggests that the demographic patterns associated with rapidly expanding populations can result in dramatic changes in genetic variation with potential relevance to fitness (Peischl and Excoffier 2015). Genomic studies are now beginning to clarify the interplay of demography, genetic variation, and adaptation, and to reveal which demographic patterns are the most relevant to adaptive potential of colonizing populations (Lohmueller 2014; Peischl and Excoffier 2015). For instance, Peischl et al. (2013) demonstrated that demographic patterns at expansion fronts of colonizing populations can result in the accumulation of deleterious mutations, and that thus “expansion load” can have dramatic long-term effects on population fitness (Peischl et al. 2013). Future empirical tests of such phenomena will doubtless be driven by application of HTS-based tools to non-model systems, including invasive populations which may offer ideal opportunities to consider the evolutionary drivers of range expansion (Rius and Darling 2014).

Many marine non-native species are associated with multiple potential vectors of introduction (“polyvectic” *sensu* Carlton and Ruiz (2015)). Not only does polyvectism potentially increase propagule pressure, it also has the capacity to increase the diversity of introduced propagule pools by sampling from different populations within the native range. Repeated introductions from multiple sources can also bring into contact different genetic lineages, and thus both promote admixture and increase the genetic diversity of introduced populations (as compared to local native populations; Rius et al. 2015). Admixture patterns have been suspected of some marine invaders, although it has proven somewhat difficult to document due to the lack of detectable genetic structure in many marine species (Rius et al. 2015). Nevertheless, growing evidence suggests an important role for admixture in driving success of introduced populations (Rius and Darling 2014). For instance, it has been suggested that admixture during range expansions could mitigate the aforementioned effects of expansion load due to gene surfing of deleterious mutations at the expansion front (Box 1; Peischl and Excoffier 2015). In 1 study, by comparing recent and past introductions of the

	Spatial contrast	Process(es)	Temporal change	Confounding factors
	Area of origin vs invaded area	Adaptation to general features of the invaded territory	Differentiation increases in time	Founder effects / admixture Mis-identification of source population(s)
	Different parts of the invaded area	Adaptation to local features of the invaded territory (eg thermal or salinity gradients)	Differentiation increases in time	Serial bottlenecks Secondary / multiple introductions
	Front vs core invasive populations	Adaptation to invasiveness Expansion load	Differentiation decreases in time (transient effect)	Mis-identification of front and core populations Difficulty to observe low-frequency deleterious mutations. Founder effects.
	Transition zone versus core	Introgression Reinforcement Capture of favorable alleles	Differentiation decreases or increases in time	Poor knowledge of genetic barriers to hybridization Coupling between physical, ecological and reproductive barriers

**Figure 2.** Different types of adaptive or phenotypic evolution accompanying an invasion, their typical effects on spatial and temporal patterns, and potential confounding factors generating similar patterns of spatial differentiation at neutral loci or at loci involved in traits not involved in adaptation to the environment.

plant *Silene vulgaris*, Keller et al. (2014) suggested that the role for admixture mainly occurs in the short term by helping introduced populations to overcome inbreeding and demographic stochasticity through an increase in heterozygosity. In marine invertebrates, heterozygosity and genetic diversity are often large even at early stages of the invasion, suggesting that this mechanism may not fully explain the success of marine NIS. The mechanisms underlying the fitness consequences of admixture still remain poorly understood, particularly in marine systems, where few sufficiently detailed empirical investigations of admixture patterns exist. In 1 outstanding example, transcriptomic analysis of the globally invasive marine crab *C. maenas* uncovered putative adaptive loci, including loci related to thermal tolerance which may play important roles in shaping genetic structure in an admixture zone between 2 different invasive lineages in the western Atlantic (Tepolt and Palumbi 2015). HTS-based methods thus may lay the groundwork for understanding how the interplay of admixture and selection shapes patterns of *C. maenas* diversity in that region, and ultimately modulates the ecological impacts of this invasive species.

The advent of HTS and its applicability to non-model organisms has opened new analytical avenues in quantifying the timing and genetic contributions of parent lineages to admixed populations, for instance, to depict the changes in the genome architecture after admixture, compare if similar changes occurred across different invasion ranges and examine fitness consequences of admixture. Whereas there is a paucity of data examining invasive populations—i.e., intra-specific level, already several studies have demonstrated the power of HTS for examining hybridization processes associated with non-native species introductions (Harrison and Larson 2014). For example, transcriptomic approaches to examine genetic architecture and consequences of historical hybridization among species of *Spartina*, a genus of frequently invasive and hybridization-prone salt marsh grasses (Ferreira de Carvalho et al. 2013). In 1

particularly ambitious study, Hohenlohe et al. (2013) used RAD sequencing to genotype over 3,000 SNPs diagnostic of the species boundary between invasive rainbow trout *Oncorhynchus mykiss* and native westslope cutthroat trout *Oncorhynchus clarkii lewisi* in the Flathead River basin of western Canada and the United States. That analysis included identification of presumably adaptive “super invasive” alleles that showed unexpectedly high admixture proportions in hybrid populations, illustrating how HTS may enable far more detailed understanding of the evolutionary consequences of admixture than have previously been possible for most organisms. Another recent study by Saarman and Pogson (2015) used SNPs derived from RAD-sequencing to detect recent hybrids and backcrosses among *Mytilus* species, ultimately concluding that their frequency was low and suggesting that introgression is not extensive between the 2 species and that ecological drivers might overcome evolutionary processes due to introgression. Investigation of introgression is potentially complicated by the need to disentangle ancient and contemporary (post-invasion) admixture events. That challenge is exemplified by the introduction of the Pacific vase tunicate *Ciona robusta* in the European native range of its congener *C. intestinalis*. Apparent footprints of contemporary hybridization were shown to be actually footprints left by an ancient secondary contact (Bouchemousse et al. 2015). Genome-wide studies allow to take into account that species boundaries are semipermeable, i.e., with gene flow which varies in intensity along the genome, and thus identify genomic regions involved in genetic incompatibilities between divergent lineages as well as to distinguish between past and contemporary admixture and hybridization processes. It is worth noting the importance of temporal sampling for exploring these relationships between propagule pressure, demographic patterns, and fitness of introduced marine population. Careful archiving of DNA, particularly in the case of rapidly shifting ranges that characterize many marine invasive species, could offer unprecedented



**Table 2.** Characteristics of studies of quantitative traits potentially subject to local adaptation in marine NIS (modified from [Tepolt 2015](#))

Species	Method	Genetic differences distinguished from plasticity?	Populations studied	Environmental variation studied <sup>a</sup>	Reproduction <sup>b</sup>
<i>Carcinus menas</i>	PD	No	NI	A	S
<i>Littorina littorea</i>	PD	No	I	A	S
<i>Crassostrea virginica</i>	CG	Yes	N	A	S
<i>Venerupis philippinarum</i>	PD	No	I	A	S
<i>Amphibalanus amphitrite</i>	CG	Yes	I	A	S
<i>Rhithropanopeus harrisi</i>	CG	Yes	NI	P, A	S
<i>Eurytemora affinis</i>	CG, EE <sup>c</sup>	Yes	NI	P, A	S
<i>Acartia tonsa</i>	PD + LC <sup>d</sup>	No or partly	N	P (1 pop), A	S
<i>Watersipora subtorquata</i>	CG	Yes	I	P, A	C
<i>Styela plicata</i>	CG	Yes	I (1 pop)	P, A	C
<i>Botrylloides violaceus</i>	PD	No	I	A	C
<i>Botrylloides spp.</i>	CG, partial RT <sup>e</sup>	Yes	I	P, A	C
<i>Botryllus schlosseri</i>	PD	No	I	A	C
<i>Bugula neritina</i>	PD	No	I	A	C
	CG	Yes	I	P, A	C
<i>Diplosoma listerianum</i>	PD	No	I	A	S
<i>Littorina saxatilis</i>	PD (cryptic species)	No	N	N	S
	PD	No	N	N	S
	PD	No	N	A	S
	CG, RT	Yes	N	P, A	S
	RT	Yes	N	A	S
	CG	Yes	N	P, A	S
	RT	No	N	A	S
<i>Asparagopsis armata</i>	CG	Yes	N	P, A	C

The first column gives the method used: quantification of phenotypic differentiation (PD) between 2 or more samples of individuals from the field; common-garden experiments (CG) and reciprocal transplants (RT). The second column indicates whether the observed phenotypic differences can be interpreted as the result of genetic variation, rather than phenotypic plasticity. In some cases CG or RT do not allow one to conclude over the genetic origin of the observed variation because individuals were transplanted to the laboratory or to other populations as adults, when they had already largely experienced the conditions of life in their original population. The third column, from [Tepolt \(2015\)](#), indicates whether native (N), introduced (I), or both types of populations were sampled. The environmental variation column indicates whether or not there was an attempt to create or compare different environmental conditions. We distinguish plasticity studies (P) in which some traits are influenced by conditions of life (individual is born in these conditions and remains there; the environment might modify his ontogeny); from accommodation studies (A) in which the short-term physiological response is the main focus. The last column describes whether the reproduction is clonal (which simplifies matters when it comes to CG or RT, but are potentially compromised by maternal or somatic effects). References are those listed in [Table 1](#) in [Tepolt \(2015\)](#).

<sup>a</sup>N (none); P (long-term plasticity); A (short term accommodation).

<sup>b</sup>Clonal (C), Sexual (S).

<sup>c</sup>EE (experimental evolution).

<sup>d</sup>LC: 1 population cultivated under laboratory conditions.

<sup>e</sup>Transplants not reciprocal.

opportunities to examine the genomic changes that accompany the various stages of invasion.

### How Common—and How Important—Is Adaptation during the Invasion Process?

Invasions at least transiently displace species and communities from their former evolutionary trajectory by putting together a species and a community with no previous coevolutionary history, or by placing a species in an environment where it has had no history of adaptation. This mismatch is bound to create strong selection pressures on both the invasive species itself and species in the recipient community ([Facon et al. 2006](#)), and hence, given appropriate genetic variance, to trigger rapid adaptive evolution. Here we focus on NIS evolution although response to NIS by natives can also be addressed ([Lau and terHorst 2015](#)). Although the premise (a mismatch between NIS and recipient environment) is simple, the specificities of

invasion contexts result in a variety of selection pressures that follow distinctive spatial and temporal patterns (see [Box 1](#)). Although population genomic tools provide novel opportunities to uncover empirical evidence of these processes and to examine the drivers of evolutionary change in invasive populations, it is nevertheless important to consider existing limitations to these approaches and to emphasize the value of integrating genomic methods with traditional approaches to studying adaptive evolution.

The simplest way to provide evidence for adaptive evolution in marine NIS is still the study of quantitative trait differentiation between populations. As noted by [Rius et al. \(2015\)](#), there is a striking paucity of studies on this subject, compared with those detailing spatial and temporal genetic structure in marine NIS. Although this situation is by no means unique to marine organisms, it is certainly more accentuated than for example in plants, in which several studies have compared introduced vs. native or successful vs. unsuccessful populations ([Colautti and Lau 2015](#)). An important point is to distinguish genetic adaptation from phenotypic plasticity; this

ideally requires full-generation culture of different populations in common garden or reciprocal transplants; otherwise, contrasts among population means for a quantitative trait might reflect effects of the environments of origin on the ontogeny rather than differences in breeding values. Unfortunately such data are very rare. In a recent review by Tepolt (2015), most of the published studies on quantitative traits that suggest adaptation in marine invasive species are in fact studies of phenotypic differentiation without common-garden or reciprocal-transplant design, most of them focusing on short-term accommodation to imposed environmental changes (e.g., temperature; see Table 2). This reflects the well-known difficulties faced by empiricists when trying to conduct inter-generational laboratory studies on marine animals. In only a few cases has a marine NIS been studied both in the native and introduced range, and using either common garden or reciprocal transplant experiments, e.g., *Eurytemora affinis* (Lee et al. 2011; Lee et al. 2012). Because of these constraints on experimental approaches, molecular tools have been used as an alternative to examine adaptation in marine NIS. They have mostly taken the form of genome scans that look for outliers of spatial differentiation among populations from both the area of origin and the area of introduction. Though we are currently aware of only 3 studies applying these techniques to marine NIS (Table 1), the approach is certain to increase in popularity given that the development of GBS considerably reduces the cost of typing several thousand of loci in individuals.

Riquet et al. (2013) looked for outliers of differentiation among all populations in a set of 15 introduced and 7 native populations of *C. fornicata*. Although they did succeed in retrieving outlier loci, nearly all of them revealed a pattern of pronounced differentiation of 1 native population from all other introduced or native populations, a pattern likely reflecting a recognized biogeographic break that occurs near Cape Canaveral in Florida. In that case, outliers of differentiation indicate genomic regions in which introgression is opposed either by some form of divergent selection or by the presence of genes coding for genetic incompatibilities between the Northern and Southern forms. No outliers were found to reveal differences between introduced and native populations or among introduced populations—providing no evidence for adaptive processes accompanying invasion.

Rohlfritsch et al. (2013) looked for  $F_{st}$  outliers in introduced populations of the oyster *C. gigas* in European coasts. Patterns of differentiation were dominated by effects of colonization history, producing distinct signatures in the northern and southern parts of the range. Interestingly,  $F_{st}$  outliers did not highlight differences between northern and southern populations, but rather tended to consistently group together a few populations from both places. The authors concluded that either parallel adaptation to some common environmental feature of these populations or the shared footprint of an independent introduction event could explain this pattern. Unfortunately, testing these hypotheses would require considerable additional sampling within the source region.

Finally, Tepolt and Palumbi (2015) studied the invasion of *C. maenas* from Europe to America with a very large number of SNPs. In sharp contrast to *Crepidula*, this study uncovered strong genetic structure in the invasion; independent introductions and serial bottlenecks have left traces in the form of high  $F_{st}$  between native and invasive populations (around 0.1) and among invasive ones (around 0–0.07), as well as significant loss in diversity. Tepolt and Palumbi (2015) did not perform a typical  $F_{st}$ -scan approach, but instead calculated the reduction in  $F_{st}$  when the 10% most differentiated loci were removed from the dataset, leaving a putatively neutral SNP panel. This selection

affects  $F_{st}$  in the native range much more than elsewhere, from which the authors conclude that local adaptation (to thermal conditions) contributes more to the overall genetic structure in the native than in the invasive range. However, this conclusion is probably hasty, as the study also revealed allelic diversity present in 1 introduced population that was absent in the native range, strongly suggesting that the true sources of the invasion remain unsampled. This finding emphasizes one of the many challenges of genomic approaches. As more loci are added, our power to detect subtle differences in genetic structure increases dramatically; however, in such conditions it becomes crucial to exactly identify source population as well as detailed invasion history (e.g., admixture events, secondary introductions, etc.). The failure to sample sources may result in a bias in background neutral expectations, which might be confounded with effects of adaptation.

It is also important to note that none of the 3 studies mentioned provided unequivocal evidence for adaptation during invasive processes, suggesting that  $F_{st}$  scans of marine NIS may face intrinsic difficulties. As noted in Section “How Do Invasions Affect Genetic Diversity in Marine Organisms?” the large effective population size of marine invertebrates might seem to favor identification of outlier loci under selection, as they may stand out even if they undergo relatively modest changes in allele frequency. However, the large effective size of many marine populations is a double-edged sword. In addition to decreasing neutral background differentiation, it also reduces the chromosomal extent of linkage disequilibria. Marker loci will only be influenced by adaptive processes occurring at loci in their close chromosomal vicinity; hence, the probability that any given marker records a particular recent selective sweep is low. In addition,  $F_{st}$  scans have intrinsic biases that are not specific to marine organisms. The most important is that they filter adaptations with a specific genetic architecture, involving few loci with very large effect, which are not necessarily the most commonly involved during invasions (Dlugosch et al. 2015); indeed adaptation depending on many genes with small effect preferentially occur by minute changes in allele frequency at each locus rather than sharp selective sweeps at a few (Le Corre and Kremer 2012). Unfortunately, life-history traits, which are likely to be under selection and usually harbor large genetic variances (Houle 1992), including in invasive populations (Facon et al. 2008), are usually affected by many loci. For this reason, an important class of adaptations may leave little trace in genome scans. Increasing genome coverage is unfortunately not a solution, and the only way to detect these evolutionary changes is through traditional measures in common garden and/or reciprocal transplant experiments.

Global  $F_{st}$  patterns, weak or strong, predominantly reflect how invasion history has shaped neutral variation. Outlier loci are in principle expected to reflect deviations from these neutral patterns. However, in practice, this inference will be valid only if the neutral expectation is correct. If relevant source populations are not included, then deviations might reflect long-established differences between the true source populations (not sampled) and the closest native population sampled, rather than recent evolution during invasion. A thorough exploration of the area of origin is therefore required. At first glance, the weak structure of many marine populations is an advantage as it should limit the impact of misidentifying the source population. However,  $F_{st}$  scans will often highlight precisely the few loci that show some significant structure in the native area. We are only starting to uncover many previously undescribed transition zones or barriers to gene flow in marine organisms; many of these barriers are visible only on small chromosomal regions associated with reproductive incompatibilities that resist introgression

during secondary contacts (Riginos and Cunningham 2007; Gosset and Bierne 2013; Fraïsse et al. 2014; Roux et al. 2014; Bouchemousse et al. 2015; Fraïsse et al. 2016). Such genomic regions can form islands of differentiation that have high chances to stand out as outlier loci—as exemplified by the *Crepidula* dataset discussed above (Riquet et al. 2013). In such cases, irrespective of how densely the genome is covered, the limiting factor is often the geographic coverage of the native range.

The overall impression left is 1 of cautious optimism. Next-generation sequencing (NGS) allows considerable increase in the number of loci, which may vastly improve the ability to detect evolutionarily relevant genetic change but is far from a panacea. Simple principles can certainly increase the efficiency of  $F_{st}$  scans. The first and most important is certainly to increase geographic coverage of the area of origin; unfortunately, typing more populations means typing more individuals. To that aim techniques such as transcriptome sequencing or RAD-seq are probably not the most cost-effective, because although they produce many loci, many of them are not polymorphic (wasting some sequencing effort) and it is hard to multiplex many individuals while maintaining reasonable coverage. A better option might be to use a first round of sequencing on a small subset of diverse individuals to define primer pairs flanking many polymorphic SNPs, and later type many individuals by amplicon sequencing, as highly multiplexed PCRs are now possible. This strategy (referred to as GT-seq for Genotyping-in-Thousands by Sequencing) is illustrated in a recent study by Campbell et al. (2015), who were able to genotype 2,068 individuals of steelhead trout at 192 loci in a single HiSeq lane. Although the number of loci remains well below that provided by RAD-seq, this technique goes in the right direction by (in principle) allowing control over how the effort is distributed between number of individuals and number of loci—and, as we have seen, there is probably more to gain by increasing the number of individuals, a conclusion also reached in by other recent population genomics studies (Benestan et al. 2015).

Still, a large fraction of the SNPs highlighted by  $F_{st}$  scans will probably resemble the *Crepidula* case, i.e., uncover previously overlooked transition zones or barriers to gene flow that divide the regional gene pool into distinct clusters, rather than recent adaptive changes in gene frequency accompanying invasions. Although this might seem disappointing, we believe that this is actually a great opportunity to gain better knowledge of the evolutionary dynamics of marine populations. In the context of marine NIS, introductions will probably create lots of novel secondary contacts (situation D in Box 1). Such situations provide the opportunity to study the interplay of adaptation, reproductive isolation, and physical obstacles to gene flow during the early phase of such contacts. This is inaccessible in the native range, where most secondary contacts are old (e.g., due to post-glacial movements). Accordingly, an hitherto unexploited possibility is to perform temporal sampling in new contact zones created by NIS and perform scans to detect outlier loci at which allele frequencies dramatically change in time (or on the contrary resist introgression in time) rather than across space.

Finally, HTS may be used to study adaptive processes in other ways than through simple  $F_{st}$  scans. First, environmental variables can be measured in natural populations and genome-wide association studies (GWAS) undertaken to extract loci whose allele frequency are particularly well correlated with them (as in, e.g., Fournier-Level et al. 2011). One pitfall of this method is that hybrid zones naturally tend to settle on sharp ecotones of environmental gradients, reflecting a coupling between different kinds of barriers to gene flow, both endogenous and environmental-driven (Bierne et al. 2011). Therefore,

as for  $F_{st}$  scans, GWAS must rely on a detailed knowledge of the geographical genetic structure and of barriers to gene flow in the native range. Finally, the challenges outlined here recommend more systematic use of common-garden or transplant designs to measure genetic differentiation at sets of candidate traits. HTS approaches, especially transcriptome studies, may here again be of great utility. Quantification of transcript copy number through sequencing or microarrays can potentially provide thousands of interesting phenotypes among which to look for signs of adaptation. The efficiency of such phenotypic scans will however depend on our ability to rigorously exclude phenotypic plasticity as a source of quantitative differences, which requires careful common-garden studies. A good example is the recent study of Sussarellu et al. (2015) revealing differential transcriptomic expression between 2 populations in a common garden experiment. A great advantage of such techniques is that they tend to relieve biases associated with the genetic architecture of traits.

## Conclusions

Twenty years of marine invasion genetic studies have illustrated the usefulness of genetic tools for addressing a wide variety of issues, ranging from identification of introduced populations to understanding the evolutionary dynamics of the invasion process. These tools have clearly revealed the extent to which cryptogenesis and cryptic cloud our assessments of non-native biogeography; have heightened our awareness of the complex relationships between propagule pressure, genetic variation, and the demographic processes of colonization; and offered early hints at the potential for introduced populations to adapt rapidly to novel environmental conditions. However, more complete understanding of these facets of marine invasion biology requires more advanced genetic tools, and the advent of HTS and associated analytical methods offers new opportunities to further explore these themes. Metabarcoding studies have already intimated at novel means of better assessing marine species distributions, with corresponding improvements in our capacity to resolve the provenance of poorly studied taxa; genomic tools for delineating species boundaries will further enhance our understanding of the degree to which species introductions have shaped marine communities. Population genomics offers compelling new approaches for understanding the role of genetic diversity in the success of introduced populations, by improving reconstructions of invasion history and better describing the genomic consequences of range expansion and intraspecific admixture. Those same tools also promise to generate new empirical evidence for rapid evolutionary change associated with marine invasions, and to provide deeper understanding of the genetic mechanisms underlying such change. All of these advanced will be buoyed by the tremendous increase in the number of loci available through HTS, as well as accessibility of entire new categories of genetic markers (e.g., neutral vs. linked to selected loci, increased use of rare variants, etc.) and increasingly sophisticated analytical tools designed to take advantage of these new data (e.g., genome mapping, Bayesian analyses). Although we have explored here the relevance of these new tools primarily to 3 central themes in marine invasion genetics, the applications are myriad. For instance, HTS may be used to address how the native species and native communities are responding to the invaders; it is equally important to turn the looking glass the other way to understand the fate of biological invasions. It is also critical to reinforce the importance of traditional tools in future research that leverages HTS. Integrated taxonomic assessments for resolving cryptic species issues, rigorously designed sampling strategies to facilitate sourcing of invasive populations, dedicated archiving of temporal DNA samples and reference data to support temporal analysis of

genetic changes throughout the invasion process, and incorporation of standard experimental quantitative trait analysis to thoroughly investigate evolutionary dynamics are all critically important to move marine invasion genetics successfully into its next phase.

## Acknowledgments

This article stems from the Doctoral Training Course “Génétique et Evolution des Organismes Marins” (Doctoral Programme SIBAGHE [University Montpellier II]) organized by François Bonhomme and Nicolas Bierne. P. D. and F. V. are thankful to them for their invitation to join this course. Meetings and workshops supported by the Marine French Connection research group GDR CNRS-Ifrermer 3445 MarCo and by the ANR grant HYSEA (ANR-12-BSV7-0011) also stimulated discussions used in this review. The United States Environmental Protection Agency, through its Office of Research and Development, supported the research described here. Though it has been subjected to Agency administrative review and approved for publication, its content does not necessarily reflect official Agency policy.

## References

- Allendorf FW, Hohenlohe PA, Luikart G, 2010. Genomics and the future of conservation genetics. *Nat Rev Genet* 11:697–709.
- Appeltans W, Ahyong Shane T, Anderson G, Angel MV, Artois T et al., 2012. The magnitude of global marine species diversity. *Curr Biol* 22:2189–2202.
- Aylagas E, Borja A, Rodriguez-Espeleta N, 2014. Environmental status assessment using DNA metabarcoding: towards a genetics based marine biotic index (gAMBI). *PLoS ONE* 9:e90529.
- Barrett RD, Schluter D, 2008. Adaptation from standing genetic variation. *Tren Ecol E* 23:38–44.
- Benestan L, Gosselin T, Perrier C, Sainte-Marie B, Rochette R et al., 2015. RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster *Homarus americanus*. *Mol Ecol* 24:3299–3315.
- Besansky NJ, Severson DW, Ferdig MT, 2003. DNA barcoding of parasites and invertebrate disease vectors: what you don't know can hurt you. *Tren Parasitol* 19:545–546.
- Bierne N, Welch J, Loire E, Bonhomme F, David P, 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol Ecol* 20:2044–2072.
- Bock DG, Caseys C, Couzens RD, Hahn MA, Heredia SM et al., 2015. What we still don't know about invasion genetics. *Mol Ecol* 24:2277–2297.
- Bouchemousse S, Haag-Liautard C, Bierne N, Viard F, 2015. Past and contemporary introgression between two strongly differentiated *Ciona* species as revealed by a post-genomic SNPs panel. *bioRxiv*. doi:10.1101/030346.
- Briski E, Ghabooli S, Bailey SA, MacIsaac HJ, 2011. Assessing invasion risk across taxa and habitats: life stage as a determinant of invasion success. *Div Dist* 17:593–602.
- Brook BW, Sodhi NS, Bradshaw CJ, 2008. Synergies among extinction drivers under global change. *Tren Ecol E* 23:453–460.
- Broquet T, Barranger A, Billard E, Bestin A, Berger R et al., 2015. The size advantage model of sex allocation in the protandrous sex-changer *Crepidula fornicata*: role of the mating system, sperm storage, and male mobility. *Am Nat* 186:404–420.
- Brunetti R, Gissi C, Pennati R, Caicci F, Gasparini F et al., 2015. Morphological evidence that the molecularly determined *Ciona intestinalis* type A and type B are different species: *Ciona robusta* and *Ciona intestinalis*. *J Zool Syst Evol Res* 53:186–193.
- Campbell NR, Harmon SA, Narum SR, 2015. Genotyping-in-thousands by sequencing (GT-seq): a cost effective SNP genotyping method based on custom amplicon sequencing. *Mol Ecol Resour* 15:855–867.
- Carlton JT, 2009. Deep invasion ecology and the assembly of communities in historical time. In: Rilov G and Crooks JA, editors. *Biological Invasions in Marine Ecosystems*. Berlin, Heidelberg: Springer, 13–56.
- Carlton JT, Ruiz GM, 2015. Anthropogenic vectors of marine and estuarine invasions: an overview framework. In: Canning-Clode J, editor. *Biological Invasions in Changing Ecosystems: Vectors, Ecological Impacts, Management and Predictions*. Warsaw/Berlin: De Gruyter Open Ltd.
- Chapman JW, Carlton JT, 1991. A test of criteria for introduced species: the global invasion by the isopod *Synidotea laevidorsalis* (Miers, 1881). *J Crustac Biol* 11:386–400.
- Colautti RI, Lau JA, 2015. Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. *Mol Ecol* 24:1999–2017.
- Comtet T, Sandionigi A, Viard F, Casiraghi M, 2015. DNA (meta)barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. *Biol Invas* 17:905–922.
- Creer S, Fonseca VG, Porazinska DL, Giblin-Davis RM, Sung W et al., 2010. Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises. *Mol Ecol* 19:4–20.
- Cristescu ME, 2014. From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. *Tren Ecol E* 29:566–571.
- Davey JL, Blaxter MW, 2011. RADSeq: next generation population genetics. *Bri Fun Genom* 9:416–423.
- Dlugosch KM, Parker IM, 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol Ecol* 17:431–449.
- Dlugosch KM, Anderson SR, Braasch J, Cang FA, Gillette HD, 2015. The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Mol Ecol* 24:2095–2111.
- Elbrecht V, Leese F, 2015. Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE* 10:e0130324.
- Ellegren H, 2014. Genome sequencing and population genomics in non-model organisms. *Tren Ecol E* 29:51–63.
- Facon B, Pointier J-P, Jarne P, Sarda V, David P, 2008. High genetic variance in life-history strategies within invasive populations by way of multiple introductions. *Curr Biol* 18:363–367.
- Facon B, Genton BJ, Shykoff J, Jarne P, Estoup A et al., 2006. A general eco-evolutionary framework for understanding bioinvasions. *Tren Ecol E* 21:130–135.
- Ferreira de Carvalho J, Poulain J, Da Silva C, Wincker P, Michon-Coudouel S et al., 2013. Transcriptome de novo assembly from next-generation sequencing and comparative analyses in the hexaploid salt marsh species *Spartina maritima* and *Spartina alterniflora* (Poaceae). *Heredity* 110:181–193.
- Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J et al., 2011. A map of local adaptation in *Arabidopsis thaliana*. *Science* 334:86–89.
- Fraïsse C, Roux C, Welch JJ, Bierne N, 2014. Gene-flow in a mosaic hybrid zone: is local introgression adaptive? *Genetics* 197:939–951.
- Fraïsse C, Belkhir K, Welch JJ, Bierne N, 2016. Local interspecies introgression is the main cause of extreme levels of intraspecific differentiation in mussels. *Mol Ecol* 21:269–286.
- Funk WC, McKay JK, Hohenlohe PA, Allendorf FW, 2012. Harnessing genomics for delineating conservation units. *Tren Ecol E* 27:489–496.
- Gagnaire P-A, Broquet T, Aurelle D, Viard F, Souissi A et al., 2015. Assessing connectivity of marine populations in the genomic era. *Evol Appl* 8:769–786.
- Geller JB, Darling JA, Carlton JT, 2010. Genetic perspectives on marine biological invasions. *Ann Rev Mar Sci* 2:401–427.
- Gollasch S, 2008. Is ballast water a major dispersal mechanism for marine organisms? In: Nentwig W, editor. *Biological Invasions*. Berlin Heidelberg: Springer.
- Gosset CC, Bierne N, 2013. Differential introgression from a sister species explains high  $F_{ST}$  outlier loci within a mussel species. *J Evol Biol* 26:14–26.
- Harrison RG, Larson EL, 2014. Hybridization, introgression, and the nature of species boundaries. *J Hered* 105 (Suppl 1):795–809.
- Hebert PDN, Ratnasingham S, deWaard JR, 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *P Roy Soc B Biol Sci* 270:S96–S99.
- Hohenlohe PA, Day MD, Amish SJ, Miller MR, Kamps-Hughes N et al., 2013. Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Mol Ecol* 22:3002–3013.



- Houle D, 1992. Comparing evolvability and variability of quantitative traits. *Genetics* 130:195–204.
- Jerde CL, Mahon AR, Chadderton WL, Lodge DM, 2011. “Sight-unseen” detection of rare aquatic species using environmental DNA. *Conserv Lett* 4:150–157.
- Katsanevakis S, Zenetos A, Belchior C, Cardoso AC, 2013. Invading European seas: assessing pathways of introduction of marine aliens. *Oce Coast Manag* 76:64–74.
- Keller SR, Fields PD, Berardi AE, Taylor DR, 2014. Recent admixture generates heterozygosity-fitness correlations during the range expansion of an invading species. *J Evol Biol* 27:616–627.
- Kelly RP, Port JA, Yamahara KM, Crowder LB, 2014. Using environmental DNA to census marine fishes in a large mesocosm. *PLoS ONE* 9:e86175.
- Knapp IS, Forsman ZH, Williams GJ, Toonen RJ, Bell JJ, 2015. Cryptic species obscure introduction pathway of the blue Caribbean sponge *Haliciona (Soestella) caerulea* (order: Haplosclerida) to Palmyra Atoll, Central Pacific. *Peer J* 3:e1170.
- Kunin V, Engelbrektson A, Ochman H, Hugenholtz P, 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* 12:118–123.
- Lau JA, terHorst CP, 2015. Causes and consequences of failed adaptation to biological invasions: the role of ecological constraints. *Mol Ecol* 24:1987–1998.
- Le Corre V, Kremer A, 2012. The genetic differentiation at quantitative trait loci under local adaptation. *Mol Ecol* 21:1548–1566.
- Lee CE, Posavi M, Charmantier G, 2012. Rapid evolution of body fluid regulation following independent invasions into freshwater habitats. *J Evol Biol* 25:625–633.
- Lee CE, Kiergaard M, Gelembiuk GW, Eads BD, Posavi M, 2011. Pumping ions: rapid parallel evolution of ionic regulation following habitat invasions. *Evolution* 65:2229–2244.
- Lindeque PK, Parry HE, Harmer RA, Somerfield PJ, Atkinson A, 2013. Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *PLoS ONE* 8:e81327.
- Lockwood JL, Cassey P, Blackburn T, 2005. The role of propagule pressure in explaining species invasions. *Tren Ecol E* 20:223–228.
- Lohmueller KE, 2014. The distribution of deleterious genetic variation in human populations. *Curr Opin Genet Dev* 29:139–146.
- McGlashan DJ, Ponniah M, Cassey P, Viard F, 2008. Clarifying marine invasions with molecular markers: an illustration based on mtDNA from mistaken calyptraeid gastropod identifications. *Biol Invas* 10:51–57.
- Mileikovsky SA, 1971. Types of larval development in marine bottom invertebrates, their distribution and ecological significance: a re-evaluation. *Mar Biol* 10:193–213.
- Molnar JL, Gamboa RL, Revenga C, Spalding MD, 2008. Assessing the global threat of invasive species to marine biodiversity. *Front Ecol Environ* 6:485–492.
- Mooney HA, Cleland EE, 2001. The evolutionary impact of invasive species. *Proc Natl Acad Sci USA* 98:5446–5451.
- Moran EV, Alexander JM, 2014. Evolutionary responses to global change: lessons from invasive species. *Ecol Lett* 17:637–649.
- Nei M, Maruyama T, Chakraborty R, 1975. The bottleneck effect and genetic variability in populations. *Evolution* 29:1–10.
- Nielsen EE, Cariani A, Aoidh EM, Maes GE, Milano I et al., 2012. Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nat Commun* 3:851.
- O’Connor TD, Fu W, Project NGENES, Group EPGaSAW, Turner E et al., 2014. Rare variation facilitates inferences of fine-scale population structure in humans. *Mol Biol E* 32:653–660.
- Pante E, Schoelink C, Puillandre N, 2015a. From integrative taxonomy to species description: one step beyond. *Syst Biol* 64:152–160.
- Pante E, Abdelkrim J, Viricel A, Gey D, France SC et al., 2015b. Use of RAD sequencing for delimiting species. *Heredity* 114:450–459.
- Pante E, Puillandre N, Viricel A, Arnaud-Haond S, Aurelle D et al., 2015c. Species are hypotheses: avoid connectivity assessment based on pillars of sand. *Mol Ecol* 24:525–544.
- Parnesan C, 2006. Ecological and evolutionary responses to recent climate change. *Ann Rev Ecol Syst* 37:637–669.
- Peischl S, Excoffier L, 2015. Expansion load: recessive mutations and the role of standing genetic variation. *Mol Ecol* 24:2084–2094.
- Peischl S, Dupanloup I, Kirkpatrick M, Excoffier L, 2013. On the accumulation of deleterious mutations during range expansions. *Mol Ecol* 22:5972–5982.
- Pfrender B, Michael E, Hawkins Charles P, Courtney GW, Creutzburg BR et al., 2010. Assessing macroinvertebrate biodiversity in freshwater ecosystems: advances and challenges in DNA-based approaches. *Q Rev Biol* 85:319–340.
- Phillips BL, Brown GP, Webb JK, Shine R, 2006. Invasion and the evolution of speed in toads. *Nature* 439:803–803.
- Pochon X, Bott NJ, Smith KF, Wood SA, 2013. Evaluating detection limits of next-generation sequencing for the surveillance and monitoring of international marine pests. *PLoS ONE* 8:e73935.
- Poloczanska ES, Brown CJ, Sydeman WJ, Kiessling W, Schoeman DS et al., 2013. Global imprint of climate change on marine life. *Nat Clim Chan* 3:919–925.
- Raupach MJ, Barco A, Steinke D, Beermann J, Laakmann S et al., 2015. The application of DNA barcodes for the identification of marine crustaceans from the North Sea and adjacent regions. *PLoS ONE* 10:e0139421.
- Riesgo A, Taboada S, Avila C, 2015. Evolutionary patterns in Antarctic marine invertebrates: an update on molecular studies. *Mari Genom* 23:1–13.
- Riginos C, Cunningham CW, 2007. Hybridization in postglacial marine habitats. *Mol Ecol* 16:3971–3972.
- Riquet F, Daguin-Thiébaud C, Ballenghien M, Bierne N, Viard F, 2013. Contrasting patterns of genome-wide polymorphism in the native and invasive range of the marine mollusc *Crepidula fornicata*. *Mol Ecol* 22:1003–1018.
- Rius M, Darling JA, 2014. How important is intraspecific genetic admixture to the success of colonising populations? *Tren Ecol Evol* 29:233–242.
- Rius M, Turon X, Bernard G, Volckaert F, Viard F, 2015. Marine invasion genetics: from spatial and temporal patterns to evolutionary outcomes. *Biol Invas* 17:869–885.
- Rohlfritsch A, Bierne N, Boudry P, Heurtebise S, Cornette F et al., 2013. Population genomics shed light on the demographic and adaptive histories of European invasion in the Pacific oyster *Crassostrea gigas*. *Evol Appl* 6:1064–1078.
- Roman J, Darling JA, 2007. Paradox lost: genetic diversity and the success of aquatic invasions. *Tren Ecol E* 22:454–464.
- Roux C, Fraise C, Castric V, Vekemans X, Pogson GH et al., 2014. Can we continue to neglect genomic variation in introgression rates when inferring the history of speciation? A case study in a *Mytilus* hybrid zone. *J Evol Biol* 27:1662–1675.
- Saarman NP, Pogson GH, 2015. Introgression between invasive and native blue mussels (genus *Mytilus*) in the central California hybrid zone. *Mol Ecol* 24:4723–4738.
- Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J et al., 2001. The population biology of invasive species. *Ann Rev Ecol Syst* 32:305–332.
- Sax DF, Brown JH, 2000. The paradox of invasion. *Glob Ecol Biogeog* 9:363–371.
- Selkoe KA, Toonen RJ, 2011. Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Mar Ecol Progr Ser* 436:291–305.
- Simberloff D, 2005. Non-native species do threaten the natural environment! *J Agr Environ Eth* 18:595–607.
- Simberloff D, 2009. The role of propagule pressure in biological invasions. *Ann Rev Ecol Syst* 40:81–102.
- Simmons M, Tucker A, Chadderton WL, Jerde CL, Mahon AR, 2016. Active and passive environmental DNA surveillance of aquatic invasive species. *Can J Fish Aquat Sci* 73:76–83.
- Sussarellu R, Huvet A, Lapegue S, Quillen V, Lelong C et al., 2015. Additive transcriptomic variation associated with reproductive traits suggest local adaptation in a recently settled population of the Pacific oyster *Crassostrea gigas*. *BMC Genomics* 16:808.
- Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E, 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol Ecol* 21:2045–2050.
- Tepolt CK, 2015. Adaptation in marine invasion: a genetic perspective. *Biol Invas* 17:887–903.

- Tepolt CK, Palumbi SR, 2015. Transcriptome sequencing reveals both neutral and adaptive genome dynamics in a marine invader. *Mol Ecol* 24:4145–4158.
- Thomsen PF, Kielgast J, Iversen LL, Møller PR, Rasmussen M et al., 2012. Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS ONE* 7:e41732.
- Treibitz AS, Hoffman JC, Grant GW, Billehus TM, Pilgrim EM, 2015. Potential for DNA-based identification of Great Lakes fauna: match and mismatch between taxa inventories and DNA barcode libraries. *Sci Rep* 5:12162.
- Viard F, Comtet T, 2016. Applications of DNA-based methods for the study of biological invasions. In: Canning-Clode J, editor. *Biological Invasions in Changing Ecosystems: Vectors, Ecological Impacts, Management and Predictions*. Berlin: De Gruyter Open, 411–435.
- Waples RS, 2015. Testing for Hardy-Weinberg proportions: have we lost the plot? *J Hered* 106:1–19.
- Weersing K, Toonen RJ, 2009. Populations genetics, larval dispersal, and connectivity in marine systems. *Mar Ecol Progr Ser* 393:1–12.
- Wilson JJ, Rougerie R, Schonfeld J, Janzen DH, Hallwachs W et al., 2011. When species matches are unavailable are DNA barcodes correctly assigned to higher taxa? An assessment using sphingid moths. *BMC Ecol* 11:18.
- Zaiko A, Samuiloviene A, Ardura A, Garcia-Vazquez E, 2015a. Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. *Mar Pollut Bull* 100:53–59.
- Zaiko A, Martinez JL, Schmidt-Petersen J, Ribicic D, Samuiloviene A et al., 2015b. Metabarcoding approach for the ballast water surveillance: an advantageous solution or an awkward challenge? *Mar Pollut Bull* 92:25–34.
- Zhan A, Xiong W, He S, MacIsaac HJ, 2014. Influence of artifact removal on rare species recovery in natural complex communities using high-throughput sequencing. *PLoS ONE* 9:e96928.
- Zhan AB, Hulak M, Sylvester F, Huang X, Adebayo AA et al., 2013. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. *Meth Ecol E* 4:558–565.

### Glossary of Terms

**Adaptation:** the selection of heritable traits that result in increased fitness.

**Admixture:** novel genetic combinations created by recombination (reproduction) between individuals with different genetic background.

**Allee effect:** when an individual of a species suffers a decrease in 1 of its fitness components at low population density. Under some threshold density, the population may risk extinction.

**Cryptic species:** a member of a pair of set of species that have been often confounded because of strong morphological similarity or lack of diagnostic morphological traits, and can usually be distinguished only by molecular markers.

**Cryptogenic species:** species which cannot be assigned to a native or non-native status because of uncertainty regarding their native range.

**DNA barcoding:** The use of short, standardized genomic loci (“DNA barcodes”) for species identification.

**Environmental DNA (eDNA):** DNA retrieved from environmental samples (sediments, water, etc.), which may contain organisms and/or short DNA molecules shed from those organisms (free, cellular debris or particle-bound).

**Founder events:** The establishment of a new population by a few colonists, resulting in a population in which only a small fraction of the genetic diversity of the source population will be found.

**Genetic drift:** random changes in frequencies of alleles in a population from one generation to another due to stochastic sampling of genetic diversity in a finite population.

**Genome scan:** Method based on the use of many loci in the hope that at least a few of them (named “outliers”) may bear the signature of selection at a nearby locus involved in adaptation.

**Genotyping By Sequencing (GBS):** massive genotyping obtained through HTS technologies; 1 commonly used GBS technique is RAD-sequencing.

**Genome wide association study (GWAS):** A form of genome scan that aims at identifying loci at which allele occurrences are highly correlated with the expression of a particular trait. By extension a genome scan that looks for loci with allele frequencies correlated with some environmental variable.

**High-throughput sequencing (HTS):** generic term used to describe a number of different recently developed sequencing technologies (e.g., Illumina, Ion torrent, Roche 454), which allow massive cost-effective sequencing; they are also known as NGS techniques.

**Introgression:** gene flow between populations or species due to repeated backcrossing from one entity to the other.

**Metabarcoding:** Characterization of complex biological communities by using HTS to barcode large numbers of individuals simultaneously in an environmental sample.

**Outlier:** a region of the genome displaying an atypical pattern of genetic differentiation indicative of direct or indirect selection processes; they are most often revealed with genome scans.

**Phenotypic plasticity:** the ability of a given genotype to display different phenotypes according to the environmental conditions.

**Population genomics:** studies of distribution of the genetic diversity at infraspecific levels by means of HTS, thus allowing genome-wide surveys of hundreds to thousands of loci.

**Propagule pressure:** combination of the number of introduction events and number of introduced individuals (propagules) to a recipient locality.

**Single nucleotide polymorphism (SNP):** a polymorphism characterized by variation at a single nucleotidic position in a short sequenced fragment. The orthology of the sequenced fragments is established either based on similarity of flanking sequences (RAD-sequencing) or amplification using specific primers (amplicon sequencing).

**Standing genetic variation:** genetic variance due to polymorphisms present in the base population, by opposition to new alleles produced by recent mutation. In the context of a recent adaptation after invasion or change in environment, adaptation based on standing variation refers to changes in frequency of alleles already present in the source population, rather than by selective sweeps of new mutations that appeared after invasion or environmental change.