

Species are hypotheses: avoid connectivity assessments based on pillars of sand

Pante Eric¹, Puillandre Nicolas², Viricel Amélia¹, Arnaud-Haond Sophie³, Aurelle Didier⁴, Castelin Magalie⁵, Chenuil Anne⁴, Destombe Christophe^{6,7}, Forcioli Didier^{8,9}, Valero Myriam^{6,7}, Viard Frédérique^{6,10}, Samadi Sarah^{2,*}

¹ Littoral, Environnement et Sociétés (LIENSs), UMR 7266 CNRS - Université de La Rochelle, 2 rue Olympe de Gouges, La Rochelle, France

² ISYEB – UMR 7205 – CNRS, MNHN, UPMC (University Paris 06), EPHE – Muséum national d'Histoire naturelle, Sorbonne Universités, Paris Cedex 05, France

³ IFREMER, UMR 212 Ecosystèmes marins Exploités, Sète, France

⁴ Aix Marseille Université, CNRS, IRD, Avignon Université, IMBE UMR 7263, Marseille, France

⁵ Aquatic Animal Health Section, Fisheries and Oceans Canada, Pacific Biological Station, Nanaimo, BC, Canada

⁶ Sorbonne Universités, UPMC, University Paris 06, Roscoff, France

⁷ CNRS, Laboratory Evolutionary Biology and Ecology of Algae, Sorbonne Universités, Université Pierre et Marie Curie (UPMC) Univ Paris 06, UMI 3614, UPMC, PUCCh, UACH, Roscoff, France

⁸ Faculté des Sciences, Université Nice-Sophia-Antipolis, Equipe Symbiose Marine UMR 7138, Nice Cedex 2, France

⁹ UMR 7138 Evolution Paris Seine, Université Pierre et Marie Curie – CNRS, 7 Quai St Bernard, France

¹⁰ Centre National de la Recherche Scientifique (CNRS), Laboratory Adaptation and Diversity in the Marine Environment, Team Diversity and Connectivity in Coastal Marine Landscapes, UMR 7144, Station Biologique de Roscoff, Roscoff, France

* Corresponding author : Sarah Samadi, fax: +33 1 40 79 38 44 ; email address : sarah.samadi@mnhn.fr

Abstract :

Connectivity among populations determines the dynamics and evolution of populations, and its assessment is essential in ecology in general and in conservation biology in particular. The robust basis of any ecological study is the accurate delimitation of evolutionary units, such as populations, metapopulations and species. Yet a disconnect still persists between the work of taxonomists describing species as working hypotheses and the use of species delimitation by molecular ecologists interested in describing patterns of gene flow. This problem is particularly acute in the marine environment where the inventory of biodiversity is relatively delayed, while for the past two decades, molecular studies have shown a high prevalence of cryptic species. In this study, we illustrate, based on marine case studies, how the failure to recognize boundaries of evolutionary-relevant unit leads to heavily biased estimates of connectivity. We review the conceptual framework within which species delimitation can be formalized as falsifiable hypotheses and show how connectivity studies can feed integrative taxonomic work and vice versa. Finally, we suggest strategies for spatial, temporal and phylogenetic sampling to reduce the probability of inadequately delimiting evolutionary units when engaging in connectivity studies.

Keywords : connectivity, marine organisms, molecular systematics, taxonomy

I. Context and problems.

Population connectivity has been the subject of an increasing number of studies for the last two decades. A Scopus search in November 2014 revealed 2,137 documents published since 1991, including 1,776 during the last 5 years. Applications of connectivity studies include: 1) identifying genetically-isolated populations of protected or endangered species that should be monitored as separate conservation units (*e.g.* Palsbøll *et al.* 2007), 2) optimizing the size and location of protected areas to create well-connected reserve networks (*e.g.* Kritzer & Sale 2004; Jones *et al.* 2007), 3) identifying stocks for fisheries management (*e.g.* Fogarty & Botsford 2007; Waples *et al.* 2008; Reiss *et al.* 2009), and 4) evaluating the potential impacts of resource exploitation on population dynamics of local communities (*e.g.* Bors *et al.* 2012; Plouviez *et al.* 2013). Besides these important applications to management and conservation, connectivity studies can also help addressing long-standing questions such as the status of marine populations as open or closed systems (*e.g.* Cowen *et al.* 2000), the identification of factors that create and maintain genetic differentiation (*e.g.* Bilton *et al.* 2002; Shank 2010), or understanding of how local adaptation can occur in high gene flow species (*e.g.* Nielsen *et al.* 2009).

In this review, we underline the paramount importance of establishing the state of taxonomic treatment and available background knowledge on the biology of organisms when engaging in connectivity assessments. These aspects are sometimes overlooked despite the fact that they should condition the sampling design and the inferences made from genetic data. One of the main considerations we develop is that properly estimating connectivity at the population level requires assessing the robustness of accepted taxonomic hypotheses. This is not trivial as for many taxa, taxonomic knowledge is inadequate or even lacking. Indeed, a great portion of the world's biodiversity remains to be described, particularly in under-explored, difficult-to-access habitats such as tropical rainforests and the marine environment (*e.g.* Appeltans *et al.* 2012). In addition, in these habitats, species descriptions are often based on few specimens, and species distribution ranges as well as ecological requirements are poorly known.

The development of DNA-sequencing techniques has considerably accelerated the rate of discovery and the documentation of species distributions. This is notably true in the marine realm, especially for poorly studied eukaryotic phyla for which few or no other characters are available as reference for taxonomic delimitations. For example, analyses of DNA characters revealed that macro- and micro-algae include many cryptic lineages that are subject to convergent morphological evolution (Leliaert *et al.* 2014). However, turning DNA-based discoveries into robustly and formally named taxonomic entities remains a

Accepted Article

difficult process (Satler *et al.* 2013). Consequently, the literature includes a large number of undescribed and un-named cryptic or pseudo-cryptic species for which identification from one study to another is challenging in the absence of formal naming and description (*e.g.* Pante *et al.* in press a). As recently underlined by Fontaine *et al.* (2012), the time lapse between specimen collection and formal species description is extremely long, averaging 21 years. A perhaps less appreciated issue is that even for abundant, commercially-important or ecologically well-studied marine taxa, species were often described long ago but are not necessarily adequately delineated (*e.g.* Uthicke *et al.* 2010; Jaafar *et al.* 2012; Mantelatto *et al.* 2014; Thomas *et al.* 2014). For instance, *Ciona intestinalis* (Linnaeus, 1767) is a model organism in evolutionary developmental biology and was among the first animals to have its genome fully sequenced (Dehal *et al.* 2002). However this taxon was recently revealed as a complex of four cryptic species (Suzuki *et al.* 2005; Caputi *et al.* 2007; Nydam & Harrison 2007; Zhan *et al.* 2010), including two widespread species that diverged ca. 3-4 Ma ago (Roux *et al.* 2013) and are not yet formally named. Knowledge gaps for marine organisms are so important (*e.g.* Knowlton 2000; Webb *et al.* 2010) that, beyond the discovery of recently-separated lineages (*i.e.* the 'emergent phylogenetic systems' defined in Carstens *et al.* 2013), taxonomic confusions at higher-rank may also impact the definition of species. For instance in 2009, Johnson *et al.* revealed that three families of deep-sea fishes artificially separated juveniles, males and females. This confusion was resolved by examining the morphology of additional specimens and by using molecular data. This result greatly changes the understanding of connectivity in terms of life history traits and ecology.

The main objective of this review is thus to stress the fact that, when conducting connectivity assessments, taxonomic status should be critically reassessed, and revised if necessary, in light of newly acquired biological data. This point is particularly crucial in marine systems, which are the focus of this review, both because the marine biota accumulates the greatest number of deep evolutionary lineages, but also because it remains vastly under-sampled (Costello *et al.* 2010; Not *et al.* 2012). We outline that the population genetics data collected for estimating connectivity may in turn be used to refine taxonomic knowledge at the species rank. Cryptic species are typically revealed by population genetic studies, phylogeographic analyses or barcoding (*e.g.* Uthicke *et al.* 2010; Jaafar *et al.* 2012; Mantelatto *et al.* 2014; Thomas *et al.* 2014). In fact, the patterns observed in connectivity studies and species delimitations belong to the same divergence continuum because they result from common ecological and micro-evolutionary processes. First, we illustrate the importance of accurate taxonomic evaluation in the analysis of connectivity by examining one specific question: the correlation between pelagic larval

duration and genetic structure estimates. We then highlight the main issues in relation with taxonomic assignment and propose a framework for preventing errors due to faulty taxonomy when conducting connectivity studies.

1. Pelagic larval duration, genetic structure, and obscure species boundaries.

The issue of poor species delimitation extends beyond individual studies, as it can affect the conclusions of meta-analyses or reviews investigating ecological trends such as the relationship between life-history traits and connectivity levels. Comparative studies including measures of genetic structure such as F_{ST} may produce biased conclusions if taxonomically-unresolved taxa are included. In *C. intestinalis* for example, the value of F_{ST} computed with microsatellites is six times higher between two populations, when each population is composed of a different cryptic species, compared to the values computed between populations within only one of these cryptic species (Zhan *et al.* 2010). If cryptic species are not treated as distinct entities, an analysis of genetic diversity among different sampling locations could lead to confound extrinsic barriers to dispersal with intrinsic reproductive barriers. We illustrate this issue by critically examining some of the source data used in a meta-analysis comparing measures of genetic differentiation across multiple species (Weersing & Toonen 2009, 2011). These meta-analyses address the question of whether early life-history traits influence connectivity and, in that case, whether this influence can be evidenced by analysis of neutral genetic markers. Different reviews and comparative studies have reached conflicting conclusions (*e.g.* Kinlan & Gaines 2003; Siegel *et al.* 2003; Shanks 2009; Weersing & Toonen 2009; Riginos *et al.* 2011; Selkoe & Toonen 2011; Dawson 2014). In their meta-analysis, Weersing & Toonen (2009) investigated how well pelagic larval duration (PLD) correlates with genetic estimates of connectivity. Their meta-analysis was based on a pseudo-random selection of papers from hits with selected keywords on electronic searches. The analyzed dataset included 130 species from 87 papers. Contrasting with the general expectation, the authors found only an extremely weak correlation between PLD and genetic structure as estimated by the overall F_{ST} . As emphasized by Selkoe & Toonen (2011), this result may be explained by various factors, including the metrics used for estimating effective dispersal, the type of markers and sampling design.

We suggest that one additional explanation for such a weak correlation may be that unrecognized cryptic species within organisms with high PLD have been included in the dataset. At least some of the selected data include questionable taxonomic entities. Among them, the reef fishes attributed to five different species of *Halichoeres* Rüppell, 1835 had among the highest F_{ST} values and unexpectedly the

highest PLD. Examination of the available data from the literature for these wrasse species shows that different taxonomic treatments have been applied for similar divergence patterns in this genus (Rocha *et al.* 2005; Rocha *et al.* 2008). The name *Halichoeres bivittatus* (Bloch, 1791) covers two divergent mitochondrial lineages whereas the names *Halichoeres brasiliensis* (Bloch, 1791) and *Halichoeres radiatus* (Linnaeus, 1758) are attributed to two mitochondrial lineages separated by a level of divergence similar to that of the lineages within *H. bivittatus*. The two mitochondrial lineages of *H. bivittatus* are allopatric, except in Bermuda where both are present, although not in the same environments. The high overall F_{ST} within *H. bivittatus* might thus result from endogenous reproductive isolation barriers rather than dispersal barriers within a metapopulation. The sharknose goby *Elacatinus evelynae* (Böhlke & Robins, 1968) is another similar example. Several mitochondrial and nuclear genetic markers and an adequate sampling of closely related species revealed three distinct and well-supported phylogenetic lineages (Taylor & Helberg 2005), but the very high F_{ST} value used in the meta-analysis by Weersing & Toonen (2009) was estimated over a sample that combined these three lineages (Taylor & Helberg 2003). These examples illustrate that depending on how the available results are interpreted in terms of taxonomic status, conclusions of comparative studies on patterns of connectivity can vary widely.

2. Theoretical framework for delimiting species: a primer.

This critical analysis underlines a crucial aspect of taxonomy: described taxa are not facts but testable hypotheses about the structure of biodiversity. This is true not only for higher taxonomic ranks, generally considered as arbitrary ranks, but also for the species rank, which is the only taxonomic category for which a “biological reality” is recognized by most scientists (see for example in Mishler 2009). The question is then how such hypotheses are formulated in reference to a theoretical background and how they are revised in light of new empirical data. The literature about the species concept is very large and the “species problem” has often been considered as unsolvable. In this debate, de Queiroz (1998) was among the first to identify that two questions of different nature are mingled. The question of defining the species category has to be distinguished from the debate on the adequacy of the criteria used in practice in the definition of the taxa ranked as species. The definition of the species category is an ontological question that may be addressed based on the theoretical ground offered by evolutionary theory (as discussed in Samadi & Barberousse 2006). How each species-taxon should be delineated is an epistemological question. It requires evaluating how well empirical data and analytic methods allow taxonomists to propose sets of organisms that fit the chosen definition of the

species category. Species-taxa are therefore considered as scientific hypotheses (species-hypotheses) engaged in a process of falsification based on the acquisition of new evidence. Although this line of thinking was only recently accepted among biologists, it closely corresponds to the working process of taxonomic revisions, which is one of the major activities of taxonomists (see discussion in Barberousse & Samadi, 2010). This analysis is also the conceptual background of the presently active field of integrative taxonomy: the multiple characters derived from various biological attributes and analyzed with various methods are used to propose stable and testable species-hypotheses. Characters are derived from biological attributes such as morphology, DNA, geographical range, habitat or behavior. Different methods are derived from the various so-called “species concepts” that are in this conceptual framework considered as criteria for species delimitations (*i.e.* biological, phylogenetic or phenetic criteria, see figure 1 for more details). This integrative framework has been thoroughly discussed; for example Camargo & Sites (2013) reviewed how characters and methods can be integrated to delimit species. As underlined by Yeates *et al.* (2011), among others, the way hypotheses are validated should not be primarily the intersection among the evidence from different datasets. Validation should rather come from the prioritization of alternative species-hypotheses based on their explanatory power as defined within the framework of evolutionary theory. In this process the quality or robustness of available species-hypotheses greatly depends on the state of knowledge about the concerned compartment of the tree of life. In fact uncertainties about available species-hypotheses may have two different causes. First, the state of the art of taxonomy may be so preliminary that different life stages of a single organism might have been erroneously classified as different species or even families (*e.g.* Johnson *et al.* 2009). Second, the evidence that two evolutionary lineages are engaged in definitive divergence may not be easy to evaluate when the time elapsed since the split between the emergent lineages is recent (the “grey zone” of de Queiroz, 1998). If the relevance of taxonomic knowledge needs to be evaluated before conducting any biological study on a given organism, we should also underline that the results of connectivity studies, if carefully conducted, can in turn provide evidence about the state of divergence among emergent lineages.

3. Where do we go from there?

Five main sources of difficulty may strongly influence measures of connectivity and taxonomic hypotheses: (1) the state of knowledge on the biology of the studied organisms, (2) the state of taxonomic treatments of the studied organisms, (3) the spatial and temporal scales of sampling, (4) the characters used to infer connectivity patterns and (5) how to synthesize information in multimarker

studies. In the following section we review these five points and give some recommendations through examples taken from diverse marine model systems. We also offer some guidelines for designing connectivity studies when little background knowledge is available.

II. Methodological considerations and guidelines for connectivity studies.

1. Background knowledge on the biology of organisms.

The availability of adequate background knowledge on the biology of organisms is crucial both for a robust taxonomic framework and for the study of connectivity. As discussed in the introduction, one of the life history traits commonly examined when studying connectivity data for marine organisms is PLD (*e.g.* Cowen & Sponaugle 2009; Shanks 2009; Weersing & Toonen 2009; Selkoe & Toonen 2011). This biological character is also informative for taxonomists when dealing with species delimitation. Other biological attributes such as reproductive strategies, sexual dimorphism or ecological specificities (*i.e.* habitat specialization, biological interactions, phenology, etc.) can significantly influence both interpretations of connectivity data and taxonomic decisions. For example, sex-specific philopatry is known to affect genetic structure (*e.g.* Engelhaupt *et al.* 2009; Baker *et al.* 2013). The existence of seasonal migrations among breeding and feeding sites, or discrete breeding sites in eels, salmon or whales (*e.g.* Bottom *et al.* 2005; Aarestrup *et al.* 2009; Baker *et al.* 2013) also has bearing on connectivity estimates. While the migratory behavior of some well-studied marine species has been characterized, the existence of seasonal migrations or discrete breeding sites is poorly known or even not considered in most marine groups. Thus, samples should be collected from different seasons when no information on migration is available. The importance of sexual dimorphism, season and location of sampling can greatly affect the observed patterns of genetic differentiation and has consequences in the evaluation of species diversity. In extreme cases, different life history stages or sexes can wrongly be attributed to different taxa (*e.g.* Johnson *et al.* 2009), confounding estimates of connectivity. When the different sexes or stages inhabit different environments such as different depths, estimating dispersal from a single life-stage or sex can bias interpretation of connectivity patterns but also taxonomic interpretations (Box 1). Other attributes, such as clonality or alternation of generations, can also complicate connectivity studies. In the case of clonal species such as seagrasses or some algae, neglecting such attributes would lead to a severe underestimation of gene flow or to overlook the presence of cryptic species (Box 2; Adjeroud *et al.* 2014). Thus, the sampling design and interpretation of genetic data in connectivity studies and taxonomic revisions should take into account life history and ecological traits that could be relevant to estimating dispersal. If such knowledge is limited or inexistent

for the species of interest, one could examine information available for closely related taxa and/or incorporate this uncertainty in the sampling design.

2. Taxonomic knowledge.

Connectivity studies generally consider described species as facts upon which the sampling scheme is established. Biogeographic range and habitat specificity are among the important attributes extracted from those species-hypotheses and are used as objective facts in the design of a connectivity study. However, an important source of uncertainty when proposing species-hypotheses is whether specimens were adequately sampled in space and time. Consequently, for organisms for which no recent taxonomic revision is available, the sampling scheme should both follow the requirements of population biology and phylogenetic systematics. Only a widespread sampling of a supposedly cosmopolitan species may reveal several diverging lineages restricted to isolated areas or distinct habitats as exemplified by the study of the supposedly ubiquitous and widespread amphipod *Eurythenes gryllus* (Lichtenstein in Mandt, 1822) (Havermans *et al.* 2013) or the copepods *Pleuromamma piseki* Farran, 1929 and *Pleuromamma gracilis* Claus, 1863 (Halbert *et al.* 2013). Collecting in under-sampled or unexplored areas may also reveal new divergent lineages. This was the case in the cosmopolitan red algae *Asparagopsis armata* Harvey, 1855 and *Asparagopsis taxiformis* (Delile) Trevisan de Saint-Léon, 1845, for which human-mediated transports increased uncertainty about taxon history and systematics (Dijoux *et al.* 2014). Sampling is also an issue for species described as geographically restricted or ecologically specialized: sampling the targeted species and closely-related taxa in other localities or habitats may help rejecting hypotheses of endemism and/or specialization. For example, recent molecular analyses revealed that chemosynthetic shrimp that were considered as distinct in cold seeps and hot vents, actually corresponded to a single species with high gene flow across the Atlantic (Box 3; Teixeira *et al.* 2013). A similar case was reported in a deep-sea mussel (genus *Idas* Jeffreys, 1876), where specimens sampled in a deep cold seep site and on vertebrate bones were previously falsely attributed to distinct species (Ritt *et al.* 2012; Thubaut *et al.* 2013). These examples illustrate the fact that the definition of the targeted group (ingroup) may be biased due to a lack of samples from specimens *a priori* identified as belonging to different species. Such sampling biases may exclude from genetic analyses a subset of populations despite its crucial role in shaping connectivity patterns.

3. Guidelines for spatial and temporal sampling of organisms.

Sampling should be guided by two objectives: adequately covering (1) both the putative geographic and ecological range of the species of interest, and (2) potentially closely-related species over a wide geographic and ecological range. Coverage at the population level should also take into account potential temporal variability in the geographic location of individuals such as differences between breeding and feeding periods, and in habitats (*e.g.* pelagic *versus* benthic).

As explained above, this two-level sampling (see Box 4) should ensure a proper estimation of the upper bound of intra-specific diversity and the lower bound of inter-specific diversity as to determine their potential overlap. The genetic distance between conspecifics and individuals from different species, known as the “barcoding gap” (Meyer & Paulay 2005), is often used as a threshold in a primary step to delimit species (Puillandre *et al.* 2012a) but its detection and interpretation highly depend on the sampling scheme. These sampling guidelines are of course difficult to follow in all situations, especially in the marine realm because of rarity, endemism, human-mediated transports, lack of knowledge on sister-species relationships for most taxa, among other factors. This strategy may point to new species-hypotheses and thus to the need of taxonomic revisions.

Taxonomic revisions are possible only if the requirements of the codes of biological nomenclature are taken into account in the sampling scheme. Indeed, an important requirement for attributing species names to redefined species hypotheses is the inclusion of type-material linked to available species-names (*e.g.* Puillandre *et al.* 2011). This last step is often lacking in molecular studies and the literature is spoiled by arbitrarily-labeled species hypotheses (*e.g.* OTUs, MOTUs, n. sp etc.) that are difficult to compare from one study to another. However, most holotypes cannot be integrated in molecular studies due to their age, their poor conservation state, or simply because they were lost. Therefore, sampling in type localities (as defined for animals in Article 76 of the International Code of Zoological Nomenclature, ICZN 1999) is a minimum requirement for naming species-hypotheses. This will both allow the corroboration or rejection of available species-hypotheses and an adequate analysis of the pattern of connectivity within the targeted species. Implementation of the aforementioned sampling strategy and interpretation of the results should be guided as far as possible by this two-level method, especially when assessing the reliability of the inferred patterns of connectivity.

4. Guidelines for the sampling of characters.

Classical markers widely used for marine organisms include genes from the phylogenetic literature and notably those mobilized in DNA barcoding initiatives. These are the *cox1* gene typically used for metazoans, (Bucklin *et al.* 2011), *RbcL* gene sequenced for macro-algae (Leliaert *et al.* 2014), but also other mitochondrial and nuclear genes such as ribosomal genes or “universal” introns (*e.g.* Jarman *et al.* 2002; Chenuil *et al.* 2010; Gérard *et al.* 2013). By using a set of unlinked genetic markers, it is possible both to refine the phylogenetic relationships and to test whether specimens from the same species are able to recombine and, conversely, whether specimens from different species cannot. This can be achieved by sequencing, for each specimen, at least two unlinked markers, such as one from the mitochondrial genome (or chloroplast genome in plants and algae) and one from the nuclear genome. It is now widely accepted that the use of a single marker, especially mitochondrial, is risky when delimiting species (Box 5), and should be used only in a first step to quickly propose primary species-hypotheses. As underlined in the review of Carstens *et al.* (2013), multi-locus approaches are required especially when considering recently diverging lineages, notably because of the genome-wide heterogeneity of introgression rates observed between such genomes (*e.g.* Roux *et al.* 2013; Destombe *et al.* 2010). Furthermore, in some taxa, classical markers sometimes lack the necessary resolution to distinguish closely related species and to infer connectivity (*e.g.* Calderón *et al.* 2006; Shearer & Coffroth 2008). Examples can be found in metazoan groups with remarkably low levels of mitochondrial evolution such as corals, when this sole genetic marker is considered (*e.g.* Shearer & Coffroth 2008; McFadden *et al.* 2011). However, in such poorly studied organisms, identifying a nuclear marker variable at the species level can be tricky (*e.g.* McFadden *et al.* 2010) and genetic similarity among individuals may reflect above-species evolutionary relationships rather than demographic connectivity (*e.g.* Miller *et al.* 2010). Clearly, these technical difficulties should be overcome thanks to single copy nuclear markers (*e.g.* Prada *et al.* 2014) and to new methods based on Next Generation Sequencing, such as the RAD-sequencing (Baird *et al.* 2008). In a single run, many independent loci covering a large part of the genome are sequenced. They are thus particularly appropriate to test for recombination among individuals and to unravel diagnostic molecular characters among taxa (*e.g.* Pante *et al.* in press b).

Even if DNA characters often provide reliable data for species delimitation, they may not always provide decisive answers. When genetic data are not fully conclusive, morphological, geographical and ecological data can help turn the scales toward alternative hypotheses. For example, Kekkonen & Hebert (2014) used the geographic criterion of whether genetic units are in sympatry or not to conclude about

the status of the detected genetic units. Our goal is not here to list all the characters and methods that can be included in the integrative taxonomy loop (as illustrated in Figure 1). We just aim to emphasize the need to discuss alternate species hypotheses, referred as (M)OTUs ((Molecular) Operational Taxonomy Units), ESU (Evolutionary Significant Units), PSH (Primary Species Hypotheses) or candidate species, depending on the method and characters, using all the available data to converge to robust, reproducible and testable hypotheses.

5. Prioritization and workflow in multi-marker approaches.

Integrative taxonomy has been proposed as a solution to overcome apparent contradictions in species-hypotheses resulting from the use of different criteria or characters (Dayrat 2005, Will *et al.* 2005). This approach is now widely accepted, although not always implemented in a fully integrated way. This results from the fact that different methodologies cannot be applied in a similar way to all taxonomic groups, since each of them has its own limitations. For example, ecological data, commonly used in species delimitation of terrestrial and shallow water organisms, are more difficult to gather for species living in areas that are not easily accessible. However, some general rules have been set up, and now, the debate does not primarily concern the definition of the species concept or the suitability of an integrative approach, but rather the relative weight and priority of the different characters and criteria used for species delimitation (DeSalle *et al.* 2005; Padial *et al.* 2010; Yeates *et al.* 2011; Camargo & Sites 2013).

The principal source of difficulty for integrative taxonomy comes from cases of recent divergence. In such cases, when taxonomy is uncertain, two alternative hypotheses explaining high differentiation values should be examined. If a species-hypothesis covers in fact two recently-separated lineages, values of the differentiation estimates between them will depend on the level of shared polymorphism. Under the hypothesis of a single species, high differentiation estimates are interpreted as population structure (low connectivity). Wang *et al.* (2008) illustrate this case by discussing how to discriminate between these two alternative hypotheses in their investigation of two morphotypes of finless porpoise (genus *Neophocaena*, Cetacea) that occur in sympatry in the strait of Taiwan. The authors apply population genetics methods to distinguish between a case of recent speciation and a case of low but recurrent gene flow. When simulating two populations with no gene flow during the last 18,000 years, simulated and observed estimates were close, both for F_{ST} and migration rate. Thus, the simulated scenario of a recent speciation event matched empirical data, leading the authors to conclude that the

two morphotypes constitute two recent species that are currently in secondary contact. In this study the two species are not reciprocally monophyletic but the population genetic approach supports the hypothesis of a complete breakage of gene flow between the two morphotypes. Reciprocal monophyly is a criterion indicating that two sets of organisms have distinct evolutionary trajectories. This phylogenetic criterion is frequently applied for species delimitation using molecular methods (*e.g.* Funk & Omland 2003) because it is much easier to evaluate than the biological criterion (inter-fecundity within species and inter-sterility among species). When the separation among studied species is ancient, the criterion of reciprocal monophyly is an operational method for detecting species. Conversely, among recently diverged species, we recommend to give advantage to population genetic methods allowing the detection of lasting breakage of gene flow over the criterion of reciprocal monophyly. New methods of analysis, mostly based on coalescent theory, allow combining both criteria and should be as much as possible used in this context (*e.g.* Yang and Rannala 2010, Leaché *et al.* 2014, Hsieh *et al.* 2014). These methodologies coupled with new genotyping techniques allow evaluating the genealogical concordance or discordance of a large set of genetic markers in a single analysis but were used only in a few pioneering studies (*e.g.* Leaché *et al.* 2014). Their applicability on poorly known organisms of the marine realm has yet to be explored, but the robustness of such analyses will greatly depend on the quality of the sampling scheme relative to the state of taxonomic and biological knowledge.

III Conclusion and implications for connectivity studies.

In many marine species, particularly those with high fecundity and large population sizes, population genetic studies have often revealed weak genetic structure. However, patterns of strong genetic structure have also been revealed and in this review, we argued that such patterns may in some cases be the outcome of incorrect sampling and/or incorrect taxonomic assignment.

We indeed showed that unreliable definition of evolutionary units can lead to either overestimation or underestimation of connectivity. This situation is analogous to type I and type II statistical errors in a hypothesis-testing framework where full connectivity is the null hypothesis. In this review we illustrated two general cases, in which both types of errors may be encountered. These two cases correspond to the number of evolutionary units being (1) larger or (2) smaller than the number of species acknowledged in the study.

In the first case, high genetic structure is equivalent to false positives and may emerge when populations from different, unrecognized or cryptic species are treated as a single species (*e.g.* Box 1). The issue lies in the interpretation of the data rather than in the data themselves, and some examples were discussed above (*e.g.* Rocha *et al.* 2005; Taylor & Helberg 2005; Rocha *et al.* 2008). The second case results in the flawed assumption of a complete lack of connectivity among taxa described as distinct species (see Box 3), preventing their recognition as potentially interdependent entities and the will to engage in connectivity studies among their populations.

As illustrated by the examples discussed in this paper, the issue of unreliable delineations of evolutionary units has been exacerbated in the past two decades with the rise of molecular data, particularly in the framework of phylogeographic studies documenting an ever-increasing number of cryptic species in the marine realm (Appeltans *et al.* 2012). In this context, we would like to reiterate that species are hypotheses that need to be revised, if necessary, in light of new information. To anticipate the possibility of false taxonomic hypotheses, we highlighted several important aspects. First, sampling design should encompass as much as possible 1) the known distribution range of the targeted organism and 2) some closely related taxa, or even some distantly related taxa in taxonomic groups where the phylogeny is poorly known (*e.g.* deep-sea organisms). The latest point is rarely considered in connectivity studies although it is an efficient way to ensure the correct delimitation of the focal species. Second, any connectivity study should be careful in its design regarding life-history traits: PLD, reproductive system characteristics (sex-ratio, clonal reproduction, etc.) and social organization, which all have important consequences on connectivity patterns but also on species delimitations. The sampling design should address the spatio-temporal stability of these life-history traits. Third, the markers used should be multiplied and diversified; in particular the sole use of mitochondrial genes should now be avoided. Altogether, the three points mentioned above are all relevant to the framework of integrative taxonomy. This approach has proven to be helpful in species delineation and meaningful when coupled with or used for connectivity studies.

Importantly, connectivity studies can reciprocally provide new data that may help defining new species in previously recognized nominal species. This is particularly well illustrated in the case of introduced or invasive marine species (see Box 6). Human-driven dispersal adds to natural dispersal, thereby broadening species ranges, sometimes at a worldwide scale. Connectivity and phylogeography studies of marine invaders have shown that important and cosmopolitan invaders are actually species complexes

(*e.g.* Bock *et al.* 2012; Zhan *et al.* 2010; Dijoux *et al.* 2014). Cryptogenic, pseudo-indigenous species and cryptic species, as defined by Carlton (2009) are new issues deserving integrative taxonomy and barcoding studies, not only to verify the accuracy of invaders reports (*e.g.* McGlashan *et al.* 2008) or to quickly report newly introduced species (Bishop *et al.* 2013), but also to examine gene flow between cryptic species that are not fully reproductively isolated. Such studies are especially important when invaders are morphologically similar to native species (*e.g.* Nydam & Harrison 2011). More generally, delineating species can be particularly challenging considering how rapidly species distribution can shift due to either introduction processes or global change. These issues increase the complexity of taxonomic assignment and biogeography studies.

Finally, poor species delimitation and the use of inadequate taxonomy assignment are not restricted to connectivity studies and have consequences on other fields of ecology and evolutionary studies, in particular experimental ecology, medicine, evolutionary developmental biology, speciation and phylogenetic studies, invasion biology research and biodiversity inventories (Bortolus 2008; Carlton 2009; Collins & Cruickshank 2013). Recent methodological and conceptual advances will contribute to avoid the pitfalls listed in this work. Thus, the emergence of integrative taxonomy and the development of new molecular techniques such as NGS-based methods will help to provide a more accurate exploration and understanding of the building-up of evolutionary units.

Acknowledgments.

The authors thank the participants and organizers of the French Marine Connectivity working group (GDR MarCo), who stimulated discussions that were the basis of this review. GDR MarCo was funded by the Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER) and the French Centre National de la Recherche Scientifique (CNRS). Salaries for E.P. and A.V. were covered by a grant to the Poitou-Charentes region (Contrat de Projet-État-Région 2007–2013), by a grant from the Fond Européen de Développement Régional (EP), and by Actions Internationales et Rayonnement of the University of La Rochelle (AV). This review benefited from the deep-sea studies developed in the French-Taiwanese project TF-DeepEvo funded by ANR and NSC (ANR 12-ISV7-0005-01) and from the project CONOTAX, funded by the French ANR (ANR-13-JSV7-0013-01).

Boxes.

Box 1. The integrative taxonomy loop: a case study of two cryptic species in the kelp *Lessonia nigrescens* Bory de Saint-Vincent, 1826 complex.

The discovery of cryptic species does not always lead to taxonomic reexamination. In many cases, cryptic species remain without taxonomic characterization, and therefore with no corresponding scientific names, long after their detection. Here we describe the case of the kelp *Lessonia nigrescens* Bory de Saint-Vincent, 1826 complex that was recently identified as two separate cryptic species thanks to a considerable sampling effort, combining phylogenetic, population genetics and ecological approaches. This kelp forms forests along the intertidal South Eastern Pacific coasts and was long considered as a single species from Cape Horn in southern Chile (ca. 55° S) to central Peru (ca. 12° S; Santelices 1989). Two cryptic species were recently identified using a combination of four markers located in the three genomic compartments - chloroplast, mitochondrion and nucleus - with 1,000 individuals covering more than 2,500 km of coastline (Tellier *et al.* 2009). The two divergent genetic lineages show a parapatric latitudinal distribution: one species extends from southern Peru (17°S) to central Chile (30° S), and the other extends from central Chile (29° S) to Chiloe Island (42°S). Between 29° S and 30° S there is a narrow area where both lineages spatially overlap in discrete patches where individuals belong to either the northern or southern species. This contact zone between 29° S and 30° S has been described as a mosaic of sites occupied by one species or the other (Tellier *et al.* 2009), rather than a true gradual transition zone, with one species replacing the other. In addition, a detailed population genetic study of gene flow in the transition zone confirmed a total absence of hybrids and mixed populations (Tellier *et al.* 2011).

Ecological differentiation between the two cryptic species was investigated, using different approaches (*e.g.* controlled response to different thermal and desiccation stresses in the lab, Oppliger *et al.* 2012; López-Cristoffanini *et al.* 2013). Both microscopic (haploid) and macroscopic (diploid) phases of the kelp life cycle were studied. Results demonstrated adaptive divergence for both phases that could explain the geographic segregation of these two cryptic species. In addition, contrasted demographic processes could be observed at various stages of their life cycle: female gametophytes from the Northern species develop rapidly but give very few oogonia per gametophyte, whereas female gametophytes from the Southern species delay their maturation, grow vegetatively and then produce numerous oogonia per individual. These results strongly advocate that, even though the origin of the genetic differentiation of

the two species is ancient, ecological and intrinsic differences are probably still important for the maintenance of reproductive isolation.

These two species were finally distinguished by subtle morphological differences leading to a formal description of two pseudo-cryptic species as *Lessonia berteroa* Montagne, 1842 and *Lessonia spicata* (Suhr) Santelices (González *et al.* 2012).

Box 2. The implications of working on organisms with clonal populations and alternation of generations: the case of seagrass meadows and algae.

Seagrass meadows were initially expected to exhibit large-scale panmixia due to the positive buoyancy of their seeds and rhizomatic fragments. These organisms are partially clonal and a blind sampling, without a proper identification of unique clonal lineages, can artificially inflate genetic structure estimates (such as F_{ST}) among meadows. When studying connectivity, the first issue is thus the ability to recognize such unique clonal lineages based on their multi-locus genotypes (Arnaud-Haond *et al.* 2007a). Even avoiding this potential bias, population genetic analyses across the distribution range of several seagrass species suggested limited gene flow (Coyer *et al.* 2004; Olsen *et al.* 2004, Arnaud-Haond *et al.* 2007b; Alberto *et al.* 2008). Fine-grained analysis of clonal lineage distribution in *Zostera marina* Linnaeus, 1753 revealed a large influence of clonality, leading to a mosaic pattern of genetic patchiness. This pattern is largely driven by patchy events of recolonization followed by the occupation of space through clonal spread, despite the likely occurrence of much larger scale effective migration (Becheler *et al.* 2010; Becheler *et al.* 2014). The genetic concept of populations, based on panmixia, remains elusive in those partially clonal organisms. Genetic differentiation may not systematically reflect lack of connectivity and thus the ecological concept of distribution continuity (*i.e.* the meadow scale) may be more accurate as a framework for connectivity analysis (Becheler *et al.* 2010).

Additionally, the general lack of isolation by distance reported across seagrass species was interpreted as rare but significant large-scale dispersal that influence the dynamics, resilience and geographic distribution of genetic diversity in seagrasses (Kendrick *et al.* 2012). This may hold for short-lived species with regular turnover and propagule recruitment, but the decomposition of the genetic distance spectrum (Rozenfeld *et al.* 2007) among meadows of the long-lived seagrass *Posidonia oceanica* (Linnaeus) Delile, 1813, shows a pattern of genetic structure likely reflecting ancient events of (re)colonization followed by long term clonal spread and accumulation of divergent mutations, rather than contemporary patterns of connectivity (Arnaud-Haond *et al.* 2014).

Last, many seaweeds alternate between a haploid (gametohyte) and a diploid (sporophyte) phase during their sexual life cycle. In such a life cycle, dominance of clonal propagation can lead to an uncoupling of haploid and diploid phases, and thus, to differences in allele frequency between gametophytes and tetrasporophytes (Sosa *et al.* 1998). This has been observed in species of red algae reproducing mainly asexually: in *Gelidium arbusculum* Bory de Saint-Vincent ex Børgesen, 1927 (Sosa *et al.* 1998) or in *Gracilaria chilensis* C.J.Bird, McLachlan & E.C.Oliveira, 1986 (Guillemin *et al.* 2008). In other species, the occurrence of sex, even occasionally, explained the lack of significant differences between haploid and diploid subpopulations: *Gelidium canariensis* (Grunow) Seoane Camba (Sosa & Garcia-Reina 1993); *Gracilaria gracilis* (Stackhouse) M.Steentoft, L.M.Irvine & W.F.Farnham 1995 (Engel *et al.* 2004); *Cladophoropsis membranacea* (Hofman Bang ex C.Agardh) Børgesen, 1905 (van der Strate *et al.* 2002) and *Chondrus crispus* Stackhouse, 1797 (Krueger-Hadfield *et al.* 2011). Genotyping both haploid and diploid individuals in these species is necessary in order to detect clonality and its effect on population structure.

Box 3: Taxonomic inflation results in the flawed *a priori* assumption of a complete lack of connectivity among morphologically distinct populations of deep-sea shrimp.

Most examples of taxonomic bias leading to distorted connectivity inferences are linked to the existence of hidden cryptic species or lineages. We here illustrate the opposite case in which morphological polymorphism led to taxonomic inflation and thus to a poor inference of connectivity patterns. This situation was found in the Alvinocaridae, a family of caridid shrimps specialized in deep chemosynthetic ecosystems. In the Atlantic, two species were described in 1988 by Williams: *A. muricola* associated with cold seeps from the western Atlantic and *A. markensis* from hydrothermal vents of the Mid Atlantic Ridge. The two species were distinguished based on a slight difference in an abdominal segment being entire or obscurely serrate. This distinction was made on juvenile specimens. The morphological distinctness of the two species was later confirmed on adults (Komai & Segonzac 2005), although on the basis of another morphological criterion: the more inflated and convex nature of the anterior part of the branchial region in *A. muricola* compared to *A. markensis*. In 2005, Komai & Segonzac cautiously identified shrimp sampled along the continental margins of eastern Africa as *A. aff. muricola*. These authors extensively discussed the pan Atlantic distribution of *A. muricola* given the puzzling absence of geographically intermediate populations. They hypothesized either the occurrence of large-scale dispersal capacities or the existence of unknown suitable sites in the intermediate area.

A barcoding approach based on the Cytochrome Oxidase I gene (*cox1*), initially aiming at retracing the phylogenetic history of alvinocarid shrimps, casted doubt on the taxonomic status of *A. muricola* and *A. markensis*, showing haplotypes shared by both species (Teixeira *et al.* 2013). Indeed, using both mitochondrial and nuclear sequences (*cox1* and 18S) together with a set of microsatellite loci, Teixeira *et al.* (2013) showed that all populations respectively attributed to these two names share a unique set of alleles at both types of genetic markers and display extremely weak genetic differentiation. This result eventually confirmed the hypothesis of intermediate populations raised by Komai & Segonzac (2005), while more importantly demonstrating that *A. markensis* and *A. muricola* should not be considered as two distinct species. Following this interpretation, the two species have to be synonymized in a single species that has a wide geographic distribution, but also a diversified range of suitable habitats in the deep sea.

The same barcoding analysis also revealed that classification at both genus and species levels should be revised within the Alvinocaridae, with possible cases of synonymy of species described as distinct genera (*e.g.* *Rimicaris hybissae* Nye, Copley & Plouviez, 2012 reported at Eastern Atlantic seeps and *Chorocaris chacei* (Williams & Rona, 1986) and at Mid-Atlantic vents). This example supports the recommendation of investigating the phylogenetic framework of the studied taxa, including closely-related taxa, before engaging into comprehensive connectivity studies.

Box 4. Large taxonomic and geographic sampling in deep-sea gastropods helps delineating species before testing hypotheses on modes of speciation.

The prevalence of allopatric *versus* sympatric modes of speciation was studied in the deep-sea Bursidae species *Bursa latitudo* Garrard, 1961, *B. fijiensis* (Watson, 1881) and *B. quirihorai* Beu, 1987 based on the biogeographic approach of Coyne & Orr (2004). These three deep-sea gastropods, inhabiting seamounts of the Norfolk and the Lord Howe ridges, and the continental slopes of New Caledonia, in the Southwest Pacific, share a largely overlapping geographic range and are not genetically structured over this area (Castelin *et al.* 2012). The question was thus to determine if speciation occurred within their present range or result from an allopatric divergence followed by a secondary contact. One requirement is to check that reproductive isolation is complete within pairs of sister species, or in other words, to check if there is a complete lack of connectivity between them. All Bursidae have a long-lived plankton-feeding larva, and most of them are widely distributed. Moreover, most are characterized by geographic rarity with sparse and highly fragmented populations. This lead to confusions in Bursidae taxonomy, with

some geographically-isolated populations attributed to distinct species names, or slightly distinct morphotypes collected at distant but not geographically-isolated sites attributed to the same species name.

The sampling strategy was thus designed to 1) extend taxonomic sampling within the Bursidae, without any *a priori* hypotheses about the systematics of the group, at both the genus and the species level; 2) extend geographic sampling to other oceans (Atlantic, Indian); and 3) extend ecological sampling by adding species from both shallow-water and deep-sea environments.

Based on this large taxon sampling, “Primary Species Hypotheses” (PSH) were drawn from shell morphology using the taxonomic literature. ESUs were then defined using the GMYC (Pons *et al.* 2006) and ABGD (Puillandre *et al.* 2012b) methods applied to *cox1* sequences and then strengthened with the phylogenetic analysis of a second mitochondrial gene (16S) and the nuclear 28S gene. All sources of evidence generated from these four steps were integrated to propose “Retained Secondary Species Hypotheses” (RSSHs).

Over the sampled geographical range, the three morphologically diagnosable entities, *B. latitudo*, *B. quirihorai* and *B. fijiensis*, are pairwise reciprocally monophyletic based on all tested loci. The dataset also revealed cryptic lineages emphasizing that species diversity in the Bursidae is underestimated, and suggesting that a taxonomic revision of the family should be performed. Among New Caledonian species, *B. fijiensis* and *B. quirihorai* are reciprocally monophyletic. These two species are the two most closely related species in the inferred phylogeny. The current biogeographic ranges of these two species and the estimated time of divergence make the scenario of sympatric speciation the most likely.

Box 5. Species delineation blurred by hybridization and suboptimal sampling: the case of *Fucus* Linnaeus, 1753.

The brown algae genus *Fucus* Linnaeus, 1753, a well-known textbook model in marine ecology, exemplifies the distribution of different morphological species in well-defined zones along a vertical intertidal gradient in the North Atlantic, from high- to low-water marks (Wahl *et al.* 2011). However, the recent use of molecular markers has deeply modified our understanding of both taxonomy and connectivity in this species complex.

Two closely related species, *Fucus spiralis* Linnaeus, 1753 and *Fucus vesiculosus* Linnaeus, 1753, are found in sympatry along eastern Atlantic rocky shores, from northern Norway to Morocco. They differ mainly in their morphology, their vertical distribution on the shore and their reproductive system. *F. spiralis* is hermaphroditic and is found in the higher intertidal, whereas *F. vesiculosus* is dioecious and is located just lower on the shore. Molecular analyses of internal transcribed spacer of nuclear ribosomal

DNA (ITS) were not efficient to separate these two taxa (Serrão *et al.* 1999) and their taxonomic status as two separate species was questioned. The use of five microsatellite loci on six populations of *F. spiralis* and ten of *F. vesiculosus*, separated by tens to hundreds of kilometers, showed that within each species individuals share diagnostic alleles whatever their geographic origin (Billard *et al.* 2005). Using gene admixture detection methods, intermediate individuals between the two species were detected along shores where the distribution of the two species overlaps at the scale of few hundred meters (Engel *et al.* 2005). The authors suggested a recent divergence of the two taxa with retention of ancestral polymorphism or introgressive hybridization with varying levels of admixture.

Interestingly, using a more refined sampling that combined transects across the zonation of the shore and quadrats at the different tidal levels, a new genetic entity previously overlooked by Engel *et al.* (2005) was detected within *F. spiralis* (Billard *et al.* 2010). Based on an integrative approach combining molecular, physiological and morphological analyses, Zardi *et al.* (2011) proposed to elevate the new genetic entity to the species level using the name *F. guiryi* G.I.Zardi, K.R.Nicastro, E.S.Serrão & G.A.Pearson, 2011. The range distribution of *F. guiryi* was shown to occur from Brittany to Morocco (Coyer *et al.* 2011). Common garden experiments performed on the three sympatric genetic taxa *F. spiralis*, *F. guiryi* and *F. vesiculosus* showed that physiological response to desiccation stress differed between species and was consistent with their respective vertical distribution on steep environmental clines in exposure time (Zardi *et al.* 2011). The opportunity of hybridization between the three species occurs only when they are physically close to each other on the shore (Zardi *et al.* 2011). In addition, it was proposed that selfing as well as limited gamete dispersal contribute to species integrity in *F. spiralis* and *F. guiryi* (Engel *et al.* 2005; Billard *et al.* 2010; Zardi *et al.* 2011). Finally, a high level of asymmetric introgression was reported (Coyer *et al.* 2011). A scenario of organelle capture via hybridization and spread of neutral nuclear alleles during range expansion was suggested to explain the recurrent occurrence of introgression in this genus. Consequently, organellar loci as barcoding tools should be used with care because of introgression.

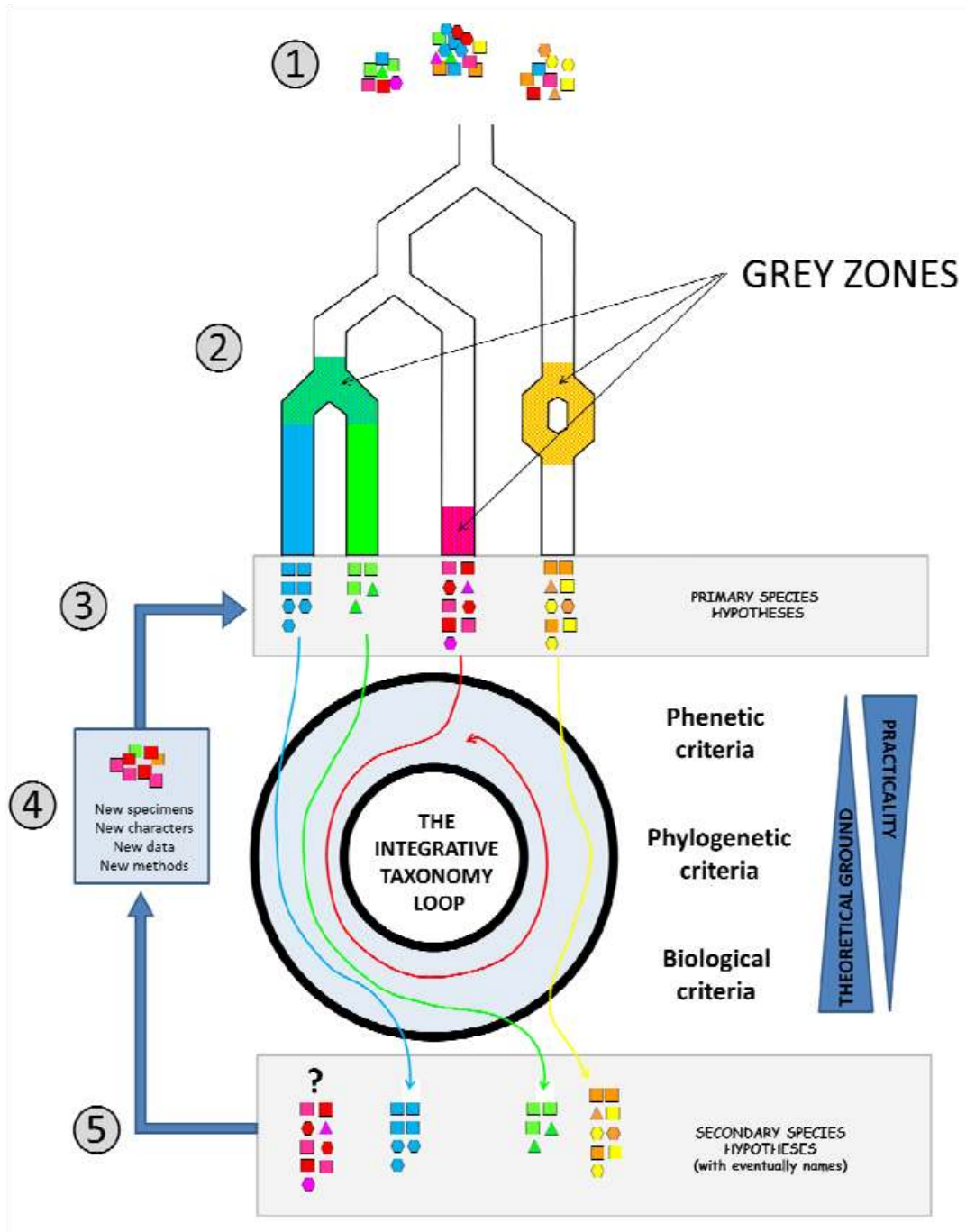
Box 6: Human-mediated transport may further increase complexity: from cryptogenic to cryptic species in marine invaders - delineating species in a changing world.

Besides natural processes of expansion and contraction of the range colonized by a species over evolutionary times, sudden species-range expansions -often through human-mediated transport- have been observed at an increasing rate since the end of the 19th century. Since the seminal work of Elton

(1958), biological invasions have increasingly received attention from the scientific community (e.g. Blackburn *et al.* 2011 and Simberloff *et al.* 2013 for recent reviews and references herein). Tightly linked to human activities particularly through shipping and aquaculture, biological invasions severely alter the biodiversity of marine ecosystems (Rilov & Crooks 2009). One of the important outcomes of marine biological invasion studies has been to document biotic homogenization at a global scale with the establishment of 'cosmopolitan' species and cryptogenic species (*i.e.* species that are neither demonstrably native nor demonstrably introduced as defined by Carlton 1996; Haydar 2012).

Interestingly, many of these cosmopolitan invasive species actually appear to be composed of morphologically-close taxa forming a species complex. This was recently exemplified in two model ascidians, the colonial tunicate *Botryllus schlosseri* (Bock *et al.* 2012) and the vase tunicate *Ciona intestinalis* (Nydam *et al.* 2011; Zhan *et al.* 2010). Both taxa are distributed worldwide. Phylogeography, phylogeny, population genetics and genomic studies showed for both taxa that they were actually characterized by marked genetic subdivision and large genealogical divergence (e.g. Roux *et al.* 2013). These two taxa are actually composed of (partially) reproductively isolated species. Some of these species came recently into secondary contact following introduction through human activities. This is the case for two members of the *Ciona intestinalis* species complex called type A and type B, which are living in sympatry after the presumably recent introduction of Type A in the Northern Atlantic. The two species are incompletely reproductively isolated in the English Channel where they are hybridizing (Nydam & Harrison 2011). The outcome of such secondary contact is unknown but may promote either adaptive introgression or speciation (Abbott *et al.* 2013). Bock *et al.* (2012) also pointed out that the different cryptic species of the *Botryllus schlosseri* complex are not characterized by similar invasive abilities, questioning the reasons for such differential abilities. Integrative taxonomy and molecular approaches are thus critical to biological invasion research. They may help identify new cryptic introduced species with different invasion abilities, understand the evolutionary history of cryptogenic species (e.g. the tunicate *Molgula manhattensis manhattensis* (De Kay, 1843), Haydar *et al.* 2011) and the fate of secondary contact between previously isolated taxa. For management purpose, they are also helpful to identify new invaders (Bishop *et al.* 2013) and to verify the accuracy of published records of invasive species (e.g. McGlashan *et al.* 2008) - especially for morphologically cryptic cosmopolitan taxa.

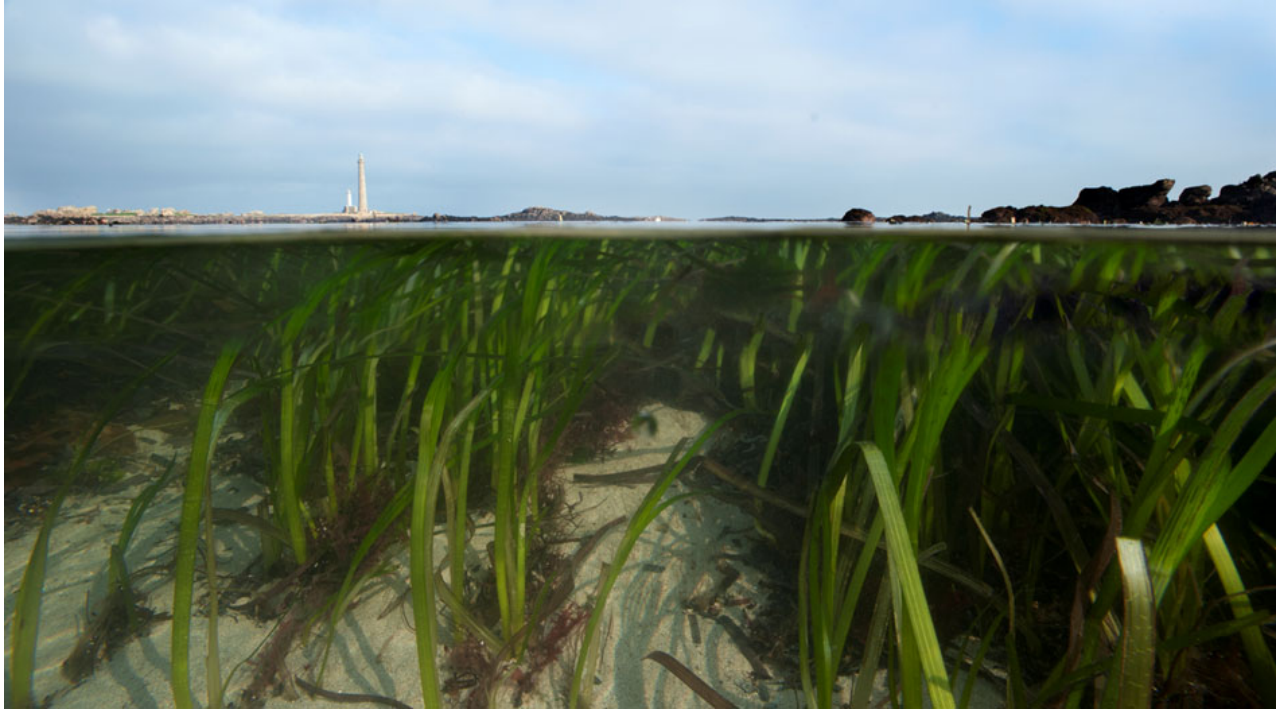
Figure1: the integrative taxonomy loop



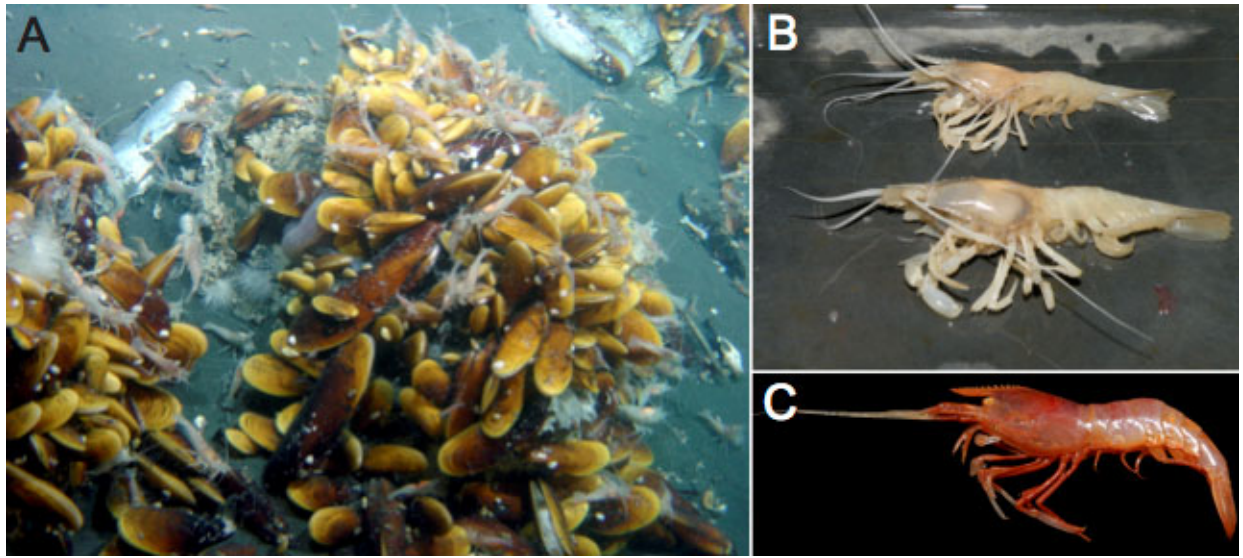
Box figure 1.



Box figure 2.



Box figure 3.



Box figure 4.



Box figure 5.



Box figure 6.



©Wilfried Thomas – Station Biologique Roscoff

Figure captions

Figure 1: The integrative taxonomy loop. Integrative taxonomy consists in analyzing different characters, with different methods, and applying different criteria of species delimitation to propose species hypotheses that are as robust as possible. The different families of criteria (as discussed in Samadi & Barberousse 2014) are listed right of the loop, with the more theoretically grounded on the bottom, and the more operational (*i.e.* easy to test practically) on the top. Within each family of criteria different kinds of characters (*i.e.* morphology, ethology, ecology, biochemical, genetic, etc.) and methods (*i.e.* distances, maximum parsimony, maximum likelihood, population genetics inferences, crossing experiments, observations, etc...) may be applied. The different steps are as follows. 1. Population and phylogenetic sampling. 2. Sampled species may be highly differentiated (blue and green), recently diverged species that are still in a “grey zone” (see Glossary) with most characters undifferentiated (pink and red), or a single species that went through a temporal split into several temporary lineages (yellow/orange). 3. Primary Species Hypotheses (PSH) are proposed, for example using morphology or a single molecular marker. 4. PSH are engaged in the integrative taxonomy loop and are evaluated, possibly with the addition of new material, using different criteria for species delimitation. The more theoretically-grounded biological criteria can be tested directly using cross experiments or indirectly with unlinked markers, and complemented with more operational criteria. 5. When possible, taxonomic decisions are taken by turning PSH into Secondary Species Hypotheses – SSH, and are named. Some lineages (*i.e.* the pink/red lineage) may stay in the loop, needing more conclusive data before being turned into SSHs. Most of the literature and methods for species delimitation focus on species that are currently in the grey zone (cf. Carstens *et al.* 2013), even though most delimitation cases fall outside of this range.

Box figure 1. The pseudo-cryptic kelp species *Lessonia spicata*. A: This species (on the left of the picture) is found in central-south Chile (29°-41°S) in areas exposed to wave actions. Sporophyte individuals can be more than 2 meters long and are anchored to the rock by a holdfast. Site Los Molles (GPS coordinates: 32°14'S-71°31'W, central Chile), photograph: Pablo Balzo. B: The gametophytes of this species are microscopic and show a sexual dimorphism with males being smaller and more branched than females. Photograph: Valeria Oppliger.

Box figure 2. *Zostera marina* meadow in Brittany. Individual shoots (ramets) are easily distinguished while rhizomatic connections linking ramets belonging to the same clonal lineage (genet) are hidden below the sandy sediment (photograph: Ifremer/Olivier Dugornay).

Box figure 3. The deep-sea shrimp *Alvinocaris muricola* and *A. markensis*. A. *Alvinocaris muricola* in situ, on a mussel bed in Western African cold seeps. (c) Ifremer, ROV Victor 6000, cruise Biozaïre. B-C. specimens of *A. muricola* (B) and *A. markensis* (c) sampled respectively on Western African cold seeps (cruise WACS, photograph Dominique Cowart) and Mid Atlantic Ridge hydrothermal vents (cruise BICOSE, photograph Laure Corbari) (c) Ifremer.

Box figure 4. The deep-sea gastropod *Bursa*. A specimen of *Bursa sp.* sampled in New Caledonia using a dredge. The other organisms sampled in the same dredge are typical of hard bottoms (e.g. stylasterids, deep-water solitary stony coral, etc...). Photograph MNHN/Pierre Lozouet.

Box figure 5. Species delineation in *Fucus*. A: Vertical distribution of furoid species on rocky shore from high to low-water mark. Site: Viana do Castelo (GPS coordinates: 41° 41' 51''N, 8° 51' 07'' W, northern Portugal), photograph: Christophe Destombe. B: Detail of reproductive structure of *Fucus guiryi* showing a typical receptacle sterile rim (arrow). Photograph: Christophe Destombe.

Box figure 6. The ascidian *Ciona*. *C. intestinalis* type A is established in the English Channel, where it shares habitats with its congener *C. intestinalis* type B. Photograph: Wilfried Thomas.

Glossary.

Biological nomenclature: set of rules that determine how names are given to taxa. In the binominal system, when a new species is described based on a given set of biological criteria applied over a set of specimens, one of them is designated as the holotype (the name-bearing specimen). When the taxonomy is revised using additional specimens and/or new biological criteria, the position of the name-bearing specimens is used to name the new taxonomic hypotheses. When studying the holotype is impossible, other categories of type material might be used (paratypes, syntypes, etc...)

Cryptic species: (see also pseudo-cryptic species): two or more species lacking obvious morphological diagnostic features that are classified as a single nominal species; it is the opposite situation compared to synonyms for which different scientific names are given to the same species.

Divergence continuum: Refers to the observation that differentiation between two emerging lineages occurs through a series of changes that accumulate during the speciation process rather than abrupt genetic changes. For a recent genomic example, see Lexer *et al.* (2014).

Grey zone: temporal zone of the genealogical network during which two lineages are definitively diverging, but the criteria used for identifying divergence might not be applicable or in agreement (see de Queiroz 1998). During this period of time, characters might differentiate at a different pace, leading to incompatible species hypotheses and justifying the need for integrative taxonomy.

Integrative taxonomy: species delimitation process in which species-hypotheses are proposed based on the integration of several lines of evidence (characters, criteria, methods) in a reproducible and falsifiable framework.

International Code of Zoological Nomenclature (ICZN): an internationally-recognized set of conventions and rules for the description of new taxa (sub-species, species and higher ranks). The ICZN follows the binomial naming system for species and sets rules to prevent conflict between names. The ICZN is independent from the International Code of Nomenclature for Algae, Fungi, and Plants (IAPT).

Population connectivity: Demographic- and evolutionary-dependence among isolated patches of individuals from a given species produced by the movement of individuals among patches.

Pseudocryptic species: morphologically overlooked species distinguished using other lines of evidence (*e.g.* DNA) that proved to be finally distinguished by subtle morphological differences (*e.g.* Box 1).

References

- Aarestrup K, Økland F, Hansen MM *et al.* (2009) Oceanic spawning migration of the European eel (*Anguilla anguilla*). *Science*, **325**, 1660.
- Abbott R, Albach D, Ansell S, *et al.* (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229-246.

- Accepted Article
- Adjeroud M, Guérécheau A, Vidal-Dupiol J, Flot J. F, Arnaud-Haond S, Bonhomme F (2014) Genetic diversity, clonality and connectivity in the scleractinian coral *Pocillopora damicornis*: a multi-scale analysis in an insular, fragmented reef system. *Marine Biology*, **161**, 531-541.
- Alberto F, Massa S, Manent P *et al.* (2008) Genetic differentiation and secondary contact zone in the seagrass *Cymodocea nodosa* across the Mediterranean-Atlantic transition region. *Journal of Biogeography*, **35**, 1279-1294.
- Appeltans W, Ahyong ST, Anderson G *et al.* (2012) The magnitude of global marine species diversity. *Current Biology*, **22**, 2189-2202.
- Arnaud-Haond S, Duarte CM, Alberto F, Serrao EA (2007a) Standardizing methods to address clonality in population studies. *Molecular Ecology*, **16**, 5115-5139.
- Arnaud-Haond S, Migliaccio M, Diaz-Almela E *et al.* (2007b) Vicariance patterns in the Mediterranean Sea: east-west cleavage and low dispersal in the endemic seagrass *Posidonia oceanica*. *Journal of Biogeography*, **34**, 963-976.
- Arnaud-Haond S, Moalic Y, Hernández-García E *et al.* (2014) Disentangling the Influence of Mutation and Migration in Clonal Seagrasses Using the Genetic Diversity Spectrum for Microsatellites. *Journal of Heredity*, **105**, 532-541.
- Baird NA, Etter PD, Atwood TS *et al.* (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE*, **3**, e3376.
- Baker CS, Steel D, Calambokidis J *et al.* (2013) Strong maternal fidelity and natal philopatry shape genetic structure in North Pacific humpback whales. *Marine Ecology Progress Series*, **494**, 291-306.
- Barberousse A, Samadi S (2010) Species from Darwin onward. *Integrative Zoology*, **5**, 187-197.
- Becheler R, Benkara E, Moalic Y, Hily C, Arnaud-Haond S (2014) Scaling of processes shaping the clonal dynamics and genetic mosaic of seagrasses through temporal genetic monitoring. *Heredity*, **112**, 114-121.
- Becheler R, Diekmann O, Hily C, Moalic Y, Arnaud-Haond S (2010) The concept of population in clonal organisms: mosaics of temporally colonized patches are forming highly diverse meadows of *Zostera marina* in Brittany. *Molecular Ecology*, **19**, 2394-2407.
- Billard E, Daguin C, Pearson G, Serrão E, Engel C, Valero M (2005) Genetic isolation between the three closely related taxa *Fucus vesiculosus*, *F. spiralis* and *F. ceranoides*. *Journal of Phycology*, **41**, 900-905.

- Billard E, Serrão E, Pearson G, Destombe C, Valero M (2010) *Fucus vesiculosus* and *spiralis* species complex: a nested model of local adaptation at the shore level. *Marine Ecology Progress Series*, **405**, 163-174.
- Bilton DT, Paula J, Bishop JDD (2002) Dispersal, genetic differentiation and speciation in estuarine organisms. *Estuarine, Coastal and Shelf Science*, **55**, 937-952.
- Bishop JDD, Roby C, Yunnice ALE *et al.* (2013) The Southern Hemisphere ascidian *Asterocarpa humilis* is unrecognised but widely established in NW France and Great Britain. *Biological Invasions*, **15**, 253-260.
- Blackburn TM, Pysek P, Bacher S, Carlton JT, Duncan RP, Jarosík V, Wilson JRU, Richardson DM (2011) A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, **26**, 333-339.
- Bock DG, Maclsaac HJ, Cristescu ME (2012) Multilocus genetic analyses differentiate between widespread and spatially restricted cryptic species in a model ascidian. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 2377-2385.
- Bors EK, Rowden AA, Maas EW, Clark MR, Shank TM (2012) Patterns of deep-sea genetic connectivity in the New Zealand region: implications for management of benthic ecosystems. *PLoS ONE*, **7**, e49474.
- Bortolus A (2008) Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *AMBIO: A Journal of the Human Environment*, **37**, 114-118.
- Bottom DL, Jones KK, Cornwell TJ, Gray A, Simenstad CA (2005) Patterns of Chinook salmon migration and residency in the Salmon River estuary (Oregon). *Estuarine Coastal and Shelf Science*, **64**, 79-93.
- Bucklin A, Steinke D, Blanco-Bercial L (2011) DNA barcoding of marine metazoa. *Annual Review of Marine Science*, **3**, 471-508.
- Calderón I, Garrabou J, Aurelle D (2006) Evaluation of the utility of COI and ITS markers as tools for population genetic studies of temperate gorgonians. *Journal of Experimental Marine Biology and Ecology*, **336**, 184-197.
- Camargo A, Sites J (2013) Species delimitation: a decade after the Renaissance. In: The species problem: ongoing issues (ed Pavlinov IY), pp 225-247. InTech – Open Access publisher.
- Caputi L, Andreakis N, Mastrotoaro F, Cirino P, Vassillo M, Sordino P (2007) Cryptic speciation in a model invertebrate chordate. *Proceedings of the National Academy of Sciences*, **104**, 9364-9369.

- Carlton J T (2009) Deep invasion ecology and the assembly of communities in historical time. In: Biological invasions in marine ecosystems (eds Rilov G, Crooks JA), pp 13-56. Springer, Berlin Heidelberg.
- Carlton JT (1996) Biological invasions and cryptogenic species. *Ecology*, **77**, 1653-1655.
- Carstens BC, Pelletier TA, Reid NM, Satler JD (2013) How to fail at species delimitation. *Molecular Ecology*, **22**, 4369-4383.
- Castelin M, Lorion J, Brisset J *et al.* (2012) Speciation patterns in gastropods with long-lived larvae from deep-sea seamounts. *Molecular ecology*, **21**, 4828-4853.
- Chenuil A, Hoareau TB, Egea E *et al.* (2010) An efficient method to find potentially universal population genetic markers, applied to metazoans. *BMC Evolutionary Biology*, **10**, 276.
- Collins RA, Cruickshank RH (2013) The seven deadly sins of DNA barcoding. *Molecular Ecology Resources*, **13**, 969-975.
- Costello MJ, Coll M, Danovaro R, Haplin P, Ojaveer H, Miloslavich P (2010) A census of marine biodiversity knowledge, resources, and future challenges. *PLoS ONE*, **5**, e11842.
- Cowen RK, Lwiza KM, Sponaugle S, Paris CB, Olson DB (2000) Connectivity of marine populations: open or closed? *Science*, **287**, 857-859.
- Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, **1**, 443-466.
- Coyer JA, Diekmann OE, Serrao EA *et al.* (2004) Population genetics of dwarf eelgrass *Zostera noltii* throughout its biogeographic range. *Marine Ecology-Progress Series*, **281**, 51-62.
- Coyer JA, Hoarau G, Costa J F *et al.* (2011). Evolution and diversification within the intertidal brown macroalgae *Fucus spiralis* / *F. vesiculosus* species complex in the North Atlantic. *Molecular Phylogenetics and Evolution*, **58**, 283-296.
- Coyne JA, Orr HA (2004) *Speciation*. Sunderland, MA: Sinauer Associates. 545 pp.
- Dayrat B (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society*, **85**, 407-415.
- Dawson MN (2014) Natural experiments and meta-analyses in comparative phylogeography. *Journal of Biogeography*, **41**, 52-65.
- de Queiroz K (1998) The general concept of species, species criteria, and the process of speciation. In: Endless forms: species and speciation (eds Howard D J, Berlocher SH), pp 57-75. Oxford University Press, Oxford.

- Dehal P, Satou Y, Campbell RK *et al.* (2002) The draft genome of *Ciona intestinalis*: Insights into chordate and vertebrate origins. *Science*, **298**, 2157-2167.
- DeSalle R, Egan MG, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**, 1905-1916.
- Destombe C, Valero M, Guillemin M-L (2010) Delineation of two sibling red algal species, *Gracilaria gracilis* and *Gracilaria dura* (Gracilariales, Rhodophyta), using multiple DNA markers: resurrection of the species *G. dura* previously described in the northern Atlantic 200 years ago. *Journal of Phycology*, **46**, 720-727.
- Dijoux L, Viard F, Payri C (2014) The more we search, the more we find: discovery of a new lineage and a new species complex in the genus *Asparagopsi*. *PLoS ONE*, **9**, e103826.
- Elton CS (1958). *The ecology of invasion by animals and plants*. 181p. London, Methuen.
- Engel C, Destombe C, Valero M (2004) Mating system and gene flow in the red seaweed *Gracilaria gracilis*: effect of haploid-diploid life history and intertidal rocky shore landscape on finescale genetic structure. *Heredity*, **92**, 289-298.
- Engel CR, Daguin C, Serrao EA (2005) Genetic entities and mating system in hermaphroditic *Fucus spiralis* and its close dioecious relative *F. vesiculosus* (Fucaceae, Phaeophyceae). *Molecular ecology*, **14**, 2033-2046.
- Engelhaupt D, Hoelzel AR, Nicholson C *et al.* (2009) Female philopatry in coastal basins and male dispersion across the North Atlantic in a highly mobile marine species, the sperm whale (*Physeter macrocephalus*). *Molecular Ecology*, **18**, 4193-4205.
- Fogarty MJ, Botsford LW (2007) Population connectivity and spatial management of marine fisheries. *Oceanography*, **20**, 112-123.
- Fontaine B, Perrard A, Bouchet P (2012) 21 years of shelf life between discovery and description of new species. *Current Biology*, **22**, R943-944.
- Funk D, Omland K (2003) Species-level paraphyly and polyphyly: Frequency, causes and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution and Systematics*, **34**, 397-423.
- Gérard K, Guilloton E, Arnaud-Haond S *et al.* (2013) PCR survey of 50 introns in animals: Cross-amplification of homologous EPIC loci in eight non-bilaterian, protostome and deuterostome phyla. *Marine Genomics*, **12**, 1-8.
- González A, Beltrán J, Hiriart - Bertrand L *et al.* (2012) Identification of cryptic species in the *Lessonia nigrescens* complex (Phaeophyceae, Laminariales). *Journal of Phycology*, **48**, 1153-1165.

- Guillemin M-L, Faugeton S, Destombe C, Viard F, Correa JA, Valero M (2008) Genetic variation in wild and cultivated populations of the haploid-diploid red alga *Gracilaria chilensis*: how farming practices favor asexual reproduction and heterozygosity. *Evolution*, **62**, 1500-1519.
- Halbert KM, Goetze E, Carlon DB (2013) High Cryptic Diversity across the Global Range of the Migratory Planktonic Copepods *Pleuromamma piseki* and *P. gracilis*. *PLoS ONE*, **8**, e77011.
- Havermans C, Sonet G, d'Acoz CDU *et al.* (2013) Genetic and morphological divergences in the cosmopolitan deep-sea amphipod *Eurythenes gryllus* reveal a diverse abyss and a bipolar species. *PLoS ONE*, **8**, e74218.
- Haydar D (2012) What is natural? The scale of cryptogenesis in the North Atlantic Ocean. *Diversity and Distributions*, **18**, 101-110.
- Haydar D, Hoarau G, Olsen JL, Stam WT, Wolff WJ (2011) Introduced or glacial relict? Phylogeography of the cryptogenic tunicate *Molgula manhattensis* (Asciacea, Pleurogona). *Diversity and Distributions*, **17**, 68-80.
- Hsieh C-H, Ko C-C, Chung C-H, Wang H-Y (2014) Multilocus approach to clarify species status and the divergence history of the Bemisia tabaci (Hemiptera: Aleyrodidae) species complex. *Molecular Phylogenetics and Evolution*, **76**, 172-80.
- ICZN (1999) International Code of Zoological Nomenclature. Fourth Edition. The International Trust for Zoological Nomenclature, London, UK. 306 pp.
- Jaafar TNAM, Taylor MI, Nor SAM, de Bruyn M, Carvalho GR (2012) DNA Barcoding Reveals Cryptic Diversity within Commercially Exploited Indo-Malay Carangidae (Teleostei: Perciformes). *PLoS ONE*, **7**, e49623.
- Jarman S, Ward R, Elliott N (2002) Oligonucleotide primers for PCR amplification of coelomate introns. *Marine Biotechnology*, **4**, 347-355.
- Johnson GD, Paxton JR, Sutton TT *et al.* (2009) Deep-sea mystery solved: astonishing larval transformations and extreme sexual dimorphisms unite three fish families. *Biology Letters*, **5**, 235-239.
- Jones GP, Srinivasan M, Almany GR (2007) Population connectivity and conservation of marine biodiversity. *Oceanography*, **20**, 100-111.
- Kekkonen M, Hebert PD (2014) DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. *Molecular ecology resources*, **14**, 706-715.
- Kendrick GA, Waycott M, Carruthers TJB *et al.* (2012) The Central Role of Dispersal in the Maintenance and Persistence of Seagrass Populations. *Bioscience*, **62**, 56-65.

- Kinlan BP, Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. *Ecology*, **84**, 2007-20.
- Knowlton N (2000) Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia*, **420**, 73–90.
- Komai T, Segonzac M (2005) A revision of the genus *Alvinocaris* Williams and Chace (Crustacea: Decapoda: Caridea: Alvinocarididae), with descriptions of a new genus and a new species of *Alvinocaris*. *Journal of Natural History*, **39**, 1111-1175.
- Kritzer JP, Sale PF (2004) Metapopulation ecology in the sea: from Levin's model to marine ecology and fisheries science. *Fish and Fisheries*, **5**, 131-140.
- Krueger-Hadfield SA, Collén J, Daguin-Thiébaud C, Valero M (2011) Genetic population structure and mating system in *Chondrus crispus* (Rhodophyta). *Journal of Phycology*, **47**, 440-450.
- Leliaert F, Verbruggen H, Vanormelingen P *et al.* (2014) DNA-based species delimitation in algae. *European Journal of Phycology*, **49**, 179-196.
- Leaché AD, Fujita MK, Minin VN, Bouckaert RR (2014). Species delimitation using genome-wide SNP data. *Systematic Biology*, **63**, 534-542.
- Lexer C, Wüest RO, Mangii S *et al.* (2014) Genomics of the divergence continuum in an African plant biodiversity hotspot, I: drivers of population divergence in *Restio capensis* (Restionaceae). *Molecular Ecology*, **23**, 4373-4386.
- López-Cristoffanini C, Tellier F, Otaíza R, Correa Juan A, Contreras-Porcía L (2013) Tolerance to air exposure: a feature driving the latitudinal distribution of two sibling kelp species. *Botanica Marina*, **56**, 431-440.
- Mantelatto FL, Pezzuto PR, Masello A, Rossi Wongtschowski CLDB, Silva Hilsdorf AW, Rossi N (2014) Molecular analysis of the commercial deep-sea crabs *Chaceon ramosae* and *Chaceon notialis* (Brachyura, Geryonidae) reveals possible cryptic species in the South Atlantic. *Deep Sea Research Part I: Oceanographic Research Papers*, **84**, 29-37.
- McFadden CS, Benayahu Y, Pante E, Thoma JN, Nevarez PA, France SC (2011) Limitations of mitochondrial gene barcoding in Octocorallia. *Molecular Ecology Resources*, **11**, 19-31.
- McFadden CS, Sánchez JA, France SC (2010) Molecular phylogenetic insights into the evolution of Octocorallia: A review. *Integrative and Comparative Biology*, **50**, 389–410.
- 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. *Biological Invasions*, **10**, 51-57.

- Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. *PLoS biology*, **3**, e422.
- Miller K, Williams A, Rowden AA, Knowles C, Dunshea G (2010) Conflicting estimates of connectivity among deep-sea coral populations. *Marine Ecology*, **31**, 144-157.
- Mishler BD (2009) Species are not uniquely real biological entities. In: Contemporary debates in philosophy of biology (eds Ayala FJ, Arp R), pp 110-122. Wiley-Blackwell, Oxford UK.
- Nielsen EE, Hemmer-Hansen J, Poulsen NA *et al.* (2009) Genomic signatures of local directional selection in a high gene flow marine organism ; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology*, **9**, 276.
- Not F, Siano R, Kooistra WH *et al.* (2012) Diversity and Ecology of Eukaryotic Marine Phytoplankton. *Advances in Botanical Research*, **64**, 1-53.
- Nydam M, Harrison R (2007) Genealogical relationships within and among shallow-water *Ciona* species (Ascidiacea). *Marine Biology*, **151**, 1839-1847.
- Nydam ML, Harrison RG (2011) Introgression despite substantial divergence in a broadcast spawning marine invertebrate. *Evolution*, **65**, 429-442.
- Olsen JL, Stam WT, Coyer JA *et al.* (2004) North Atlantic phylogeography and large-scale population differentiation of the seagrass *Zostera marina* L. *Molecular Ecology*, **13**, 1923-1941.
- Oppliger LV, Correa JA, Engelen AH *et al.* (2012) Temperature Effects on Gametophyte Life-History Traits and Geographic Distribution of Two Cryptic Kelp Species. *PLoS ONE*, **7**, e39289.
- Padial JM, Miralles A, De la Riva I, Vences M (2010) Review: The integrative future of taxonomy. *Frontiers in Zoology*, **7**, 1-14.
- Palsbøll PJ, Bérubé M, Allendorf F (2007) Identification of management units using population genetic data. *Trends in Ecology and Evolution*, **22**, 11-16.
- Pante E, Schoelincq C, Puillandre N (in press a) From integrative taxonomy to species description: one step beyond. *Systematic Biology* doi: 10.1093/sysbio/syu083
- Pante E, Abdelkrim J, Viricel A, Gey D, France S, Boisselier MC, Samadi S (in press b) Use of RAD sequencing for delimiting species. *Heredity* doi:10.1038/hdy.2014.105
- Plouviez S, Schultz T, McGinnis G, Minshall H, Rudder M, Van Dover CL (2013) Genetic diversity of hydrothermal-vent barnacles in Manus Basin. *Deep Sea Research Part I: Oceanographic Research Papers*, **82**, 73-79.
- Pons J, Barraclough TG, Gomez-Zurita J, *et al.* (2006) Sequence based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, **55**, 595–609.

- Prada C, McIlroy SE, Beltrán DM, Valint DJ, Ford SA, Hellberg ME, Coffroth MA (2014) Cryptic diversity hides host and habitat specialization in a gorgonian-algal symbiosis. *Molecular Ecology*, **23**, 3330–3340.
- Puillandre N, Lambert A, Brouillet S, Achaz G (2012b) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, **21**, 1864–1877.
- Puillandre N, Macpherson E, Lambourdière J, Cruaud C, Boisselier-Dubayle MC, Samadi S (2011) Barcoding type specimens helps to identify synonyms and an unnamed new species in *Eumunida* Smith, 1883 (Decapoda: Eumunidae). *Invertebrate Systematics*, **25**, 322–333.
- Puillandre N, Modica MV, Zhang Y *et al.* (2012a) Large-scale species delimitation method for hyperdiverse groups. *Molecular ecology*, **21**, 2671–2691.
- Reiss H, Hoarau G, Dickey-Collas M, Wolff WJ (2009) Genetic population structure of marine fish: mismatch between biological and fisheries management units. *Fish and Fisheries*, **10**, 361–395.
- Riginos C, Douglas KE, Jin Y, Shanahan DF, Trembl EA (2011) Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, **34**, 566–75.
- Rilov G, Crooks Eds. (2009) *Biological Invasions in Marine Ecosystems: Ecological, Management, and Geographic Perspectives*. Ecological Studies. 641 pages. Springer.
- Ritt B, Duperron S, Lorion J, Sara Lazar C, Sarrazin J (2012) Integrative study of a new cold-seep mussel (Mollusca: Bivalvia) associated with chemosynthetic symbionts in the Marmara Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, **67**, 121–132.
- Rocha LA, Robertson DR, Roman J, Bowen BW (2005) Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society B*, **272**, 573–579.
- Rocha LA, Rocha CR, Robertson DR, Bowen BW (2008) Comparative phylogeography of Atlantic reef fishes indicates both origin and accumulation of diversity in the Caribbean. *BMC Evolutionary Biology*, **8**, 157.
- Roux C, Tasgkogeorga G, Bierne N, Galtier N (2013) Crossing the Species Barrier: Genomic Hotspots of Introgression between Two Highly Divergent *Ciona intestinalis* Species. *Molecular Biology and Evolution*, **30**, 1574–1587.
- Rozenfeld AF, Arnaud-Haond S, Hernández-García E *et al.* (2007) Spectrum of genetic diversity and networks of clonal populations. *Journal of the Royal Society Interface*, **4**, 1093–1102.
- Samadi S, Barberousse A (2006) The tree, the network, and the species. *Biological Journal of the Linnean Society*, **89**, 509–521.

- Accepted Article
- Samadi S, Barberousse A. (2014). Species. *In press*. In: Handbook of Evolution Thinking in the Sciences (Heams, Huneman, Lecointre, Silberstein, eds). Springer.
- Santelices B (1989) *Algas marinas de Chile*. 399 pp. Ediciones Universidad Católica de Chile, Santiago, Chile.
- Satler JD, Carstens BC, Hedin M (2013) Multilocus Species Delimitation in a Complex of Morphologically Conserved Trapdoor Spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*). *Systematic Biology* **62**, 805-823.
- Selkoe KA, Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series*, **436**, 291-305 .
- Serrão EA, Alice LA, Brawley SH (1999) Evolution of the Fucaceae (Phaeophyceae) inferred from nrDNA-ITS. *Journal of Phycology*, **35**, 382-394.
- Shank TM (2010) Seamounts: Deep-oceans laboratories of faunal connectivity, evolution, and endemism. *Oceanography*, **23**, 108-122.
- Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. *Biological Bulletin*, **216**, 373-85.
- Shearer TL, Coffroth MA (2008) Barcoding corals: limited by interspecific divergence, not intraspecific variation. *Molecular Ecology Resources*, **8**, 247–255.
- Siegel DA, Kinlan BP, Gaylord B, Gaines SD (2003) Lagrangian descriptions of marine larval dispersion. *Marine Ecology Progress Series*, **260**, 83-96.
- Simberloff D, Martin J-L, Genovesi P, Maris V, Wardle DA, Aronson J, Courchamp F, Galil B, García-Berthou E, Pascal M, Pyšek P, Sousa R, Tabacchi E, Vilà M (2013) Impacts of biological invasions: what's what and the way forward. *Trends in Ecology & Evolution*, **28**, 58-66.
- Sosa P, Valero M, Batista F, Gonzalez-Pérez M (1998) Genetic structure of natural populations of *Gelidium* species: A re-evaluation of results. *Journal of Applied Phycology*, **10**, 279-284.
- Sosa PA, Garcia-Reina G (1993) Genetic variability of *Gelidium canariensis* (Rhodophyta) determined by isozyme electrophoresis. *Journal of Phycology*, **29**, 118-124.
- Suzuki MM, Nishikawa T, Bird A (2005) Genomic approaches reveal unexpected genetic divergence within *Ciona intestinalis*. *Journal of Molecular Evolution*, **61**, 627-635.
- Taylor MS, Hellberg ME (2003) Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. *Science*, **299**, 107-109.
- Taylor MS, Hellberg ME (2005) Marine radiations at small geographical scales: speciation in Neotropical reef gobies (*Elacatinus*). *Evolution*, **59**, 374-385.

- Teixeira S, Olu K, Decker C *et al.* (2013) High connectivity across the fragmented chemosynthetic ecosystems of the deep Atlantic Equatorial Belt: efficient dispersal mechanisms or questionable endemism? *Molecular Ecology*, **22**, 4663-4680.
- Tellier F, Meynard A, Correa J, Faugeron S, Valero M (2009) Phylogeographic analyses of the 30 degrees S south-east Pacific biogeographic transition zone establish the occurrence of a sharp genetic discontinuity in the kelp *Lessonia nigrescens*: Vicariance or parapatry? *Molecular Phylogenetics and Evolution*, **53**, 679-693.
- Tellier F, Tapia J, Faugeron S, Destombe C, Valero M (2011) The *Lessonia nigrescens* species complex (Laminariales, Phaeophyceae) shows strict parapatry and complete reproductive isolation in a secondary contact zone. *Journal of Phycology*, **47**, 894-903.
- Thomas Jr RC, Willette DA, Carpenter KE, Santos MD (2014) Hidden Diversity in Sardines: Genetic and Morphological Evidence for Cryptic Species in the Goldstripe Sardinella, *Sardinella gibbosa* (Bleeker, 1849). *PLoS ONE*, **9**, e84719.
- Thubaut J, Puillandre N, Faure B, Cruaud C, Samadi S (2013) The contrasted evolutionary fates of deep-sea chemosynthetic mussels (Bivalvia, Bathymodiolinae). *Ecology and evolution*, **3**, 4748-4766.
- Uthicke S, Byrne M, Conand C (2010) Genetic barcoding of commercial Bêche de mer species (Echinodermata: Holothuroidea). *Molecular Ecology Resources*, **10**, 634-646.
- van der Strate HJ, van de Zande L, Stam WT, Olsen JL (2002) The contribution of haploids, diploids and clones to fine-scale population structure in the seaweed *Cladophoropsis membranacea* (Chlorophyta). *Molecular Ecology*, **11**, 329-345.
- Wahl M, Jormalainen V, Eriksson BK *et al.* (2011) Stress ecology in *Fucus*: abiotic, biotic and genetic interactions. *Advances in marine biology*, **59**, 37-106.
- Wang JY, Frasier TR, Yang SC, White BN (2008) Detecting recent speciation events: the case of the finless porpoise (genus *Neophocaena*). *Heredity*, **101**, 145-155.
- Waples RS, Punt AE, Cope JM (2008) Integrating genetic data into management of marine resources: how can we do it better?. *Fish and Fisheries*, **9**, 423-449.
- Webb TJ, Berghe EV, O'Dor R (2010) Biodiversity's big wet secret: the global distribution of marine biological records reveals chronic under-exploration of the deep pelagic ocean. *PLoS ONE*, **5**, e10223.
- Weersing K, Toonen RJ (2009) Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, **393**, 12.

Will KW, Mishler BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative taxonomy. *Systematic biology*, **54**, 844-851.

Williams AB (1988) New marine decapod crustaceans from waters influenced by hydrothermal, discharge, brine and hydrocarbon seepage. *Fishery Bulletin*, **86**, 263–287.

Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci U S A*, **107**, 9264–9.

Yeates DK, Seago A, Nelson L, Cameron SL, Joseph L, Trueman JW (2011) Integrative taxonomy, or iterative taxonomy?. *Systematic Entomology*, **36**, 209-217.

Zardi GI, Nicastro KR, Canovas F, Costa JF, Serrao EA, Pearson GA (2011) Adaptive Traits Are Maintained on Steep Selective Gradients despite Gene Flow and Hybridization in the Intertidal Zone. *PLoS ONE*, **6**, e19402.

Zhan A, Maclsaac HJ, Cristescu ME (2010) Invasion genetics of the *Ciona intestinalis* species complex: from regional endemism to global homogeneity. *Molecular Ecology*, **19**, 4678-4694.