



**HAL**  
open science

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

Eric Pante, Nicolas Puillandre, Amélia Viricel, Sophie Arnaud-Haond, Didier Aurelle, Magalie Castelin, Anne Chenuil, Christophe Destombe, Didier Forcioli, Myriam Valero, et al.

### ► To cite this version:

Eric Pante, Nicolas Puillandre, Amélia Viricel, Sophie Arnaud-Haond, Didier Aurelle, et al.. Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Molecular Ecology*, 2015, 24 (3), pp.525-544. hal-02002440

**HAL Id: hal-02002440**

**<https://hal.science/hal-02002440v1>**

Submitted on 31 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MOLECULAR ECOLOGY

**Species are hypotheses: avoid basing connectivity assessments on pillars of sand.**

Journal:	<i>Molecular Ecology</i>
Manuscript ID:	Draft
Manuscript Type:	Invited Reviews and Syntheses
Date Submitted by the Author:	n/a
Complete List of Authors:	Pante, Eric; UMR 7266 CNRS - Université de La Rochelle, Puillandre, Nicolas; MNHN, Systematique & Evolution Viricel, Amélia; UMR 7266 CNRS - Université de La Rochelle, Arnaud-Haond, Sophie; IFREMER, Aurelle, Didier; CNRS UMR 6540 – DIMAR, Station marine d'Endoume Castelin, Magalie; Fisheries and Oceans Canada, Chenuil, Anne; CNRS UMR 6540 – DIMAR, Station marine d'Endoume Destombe, Christophe; CNRS UPMC, Forcioli, Didier; University of Nice Sophia Antipolis, , Equipe Symbiose Marine UMR 7138 Valero, Myriam; CNRS UPMC, Viard, Frederique; CNRS UPMC, Samadi, Sarah; MNHN, Systematique & Evolution
Keywords:	Connectivity, Marine organisms, Molecular systematics, Taxonomy

# 1 Species are hypotheses: avoid basing connectivity assessments on pillars of sand

2

3 Eric Pante\*<sup>1</sup>, Nicolas Puillandre\*<sup>2</sup>, Amélia Viricel\*<sup>1</sup>, Sophie Arnaud-Haond<sup>3</sup>, Didier Aurelle<sup>4</sup>, Magalie  
4 Castelin<sup>5</sup>, Anne Chenuil<sup>4</sup>, Christophe Destombe<sup>6-7</sup>, Didier Forcioli<sup>8-9</sup>, Myriam Valero<sup>6-7</sup>, Frédérique Viard<sup>6-10</sup>,  
5 Sarah Samadi<sup>2</sup>

6

7 \*authors with equal contributions.

8

9 <sup>1</sup> Littoral, Environnement et Sociétés (LIENSs), UMR 7266 CNRS - Université de La Rochelle, 2 rue Olympe de Gouges,  
10 17042 La Rochelle, France

11 <sup>2</sup> ISYEB – UMR 7205 – CNRS, MNHN, UPMC (University Paris 06), EPHE – Muséum national d’Histoire naturelle,  
12 Sorbonne Universités, CP26, 57 rue Cuvier, F-75231 Paris cedex 05, France

13 <sup>3</sup> IFREMER, UMR 212 Ecosystèmes marins Exploités, F-34203 Sète, France

14 <sup>4</sup> Aix Marseille Université, CNRS, IRD, Avignon Université, IMBE UMR 7263, 13397, Marseille, France

15 <sup>5</sup> Aquatic Animal Health Section, Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road,  
16 Nanaimo, BC, Canada V9T 6N7

17 <sup>6</sup> Sorbonne Universités, UPMC, University Paris 06, Station Biologique de Roscoff, F-29680 Roscoff, France

18 <sup>7</sup> CNRS, Laboratory Evolutionary Biology and Ecology of Algae, , Sorbonne Universités, Université Pierre et Marie Curie  
19 (UPMC) Univ Paris 06, UMI 3614, UPMC, PUCCh, UACH, Station Biologique de Roscoff, F-29680 Roscoff, France

20 <sup>8</sup> Université Nice-Sophia-Antipolis, Equipe Symbiose Marine UMR 7138, Faculté des Sciences, Parc Valrose, 06108 Nice  
21 cedex 2, France

22 <sup>9</sup> Université Pierre et Marie Curie – CNRS, UMR 7138 Evolution Paris Seine, 7 Quai St Bernard, 75252 Paris Cedex 05,  
23 France

24 <sup>10</sup> CNRS, Laboratory Adaptation and Diversity in the Marine Environment, Team Diversity and Connectivity in Coastal  
25 Marine Landscapes, UMR 7144, Station Biologique de Roscoff, F-29680 Roscoff, France

26 Centre National de la Recherche Scientifique (CNRS), UMR 7144, Divco team, Station Biologique de Roscoff, Roscoff,  
27 France

28

29 Corresponding author: Sarah Samadi

30 Pr Sarah Samadi

31 Muséum National d’Histoire Naturelle, Département Systématique et Evolution

32 Institut de Systématique, Evolution, Biodiversité, UMR 7205 CNRS MNHN UPMC EPHE

33 57 rue Cuvier CP 26

34 75231 Paris Cedex 05, France

35 Phone: 33 1 40 79 37 59

36 E-mail: sarah@mnhn.fr

37

38 **Abstract.**

39

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

54

55

56

57 **I. Context and problems.**

58

59 Population connectivity, involving the demographic and/or evolutionary interdependence of populations

60 through individual movements across their species range, has been subject to an increasing number of

61 studies for the last two decades (Scopus search for “population connectivity”, accessed on August 23<sup>th</sup>,

62 2014: 2,060 documents published since 1993, including 1,700 during the last 5 years). Connectivity studies

63 have bearing on both applied and theoretical research. For instance, assessing the extent of connectivity

64 among populations has become a key aspect in diverse areas of conservation biology. Applications of

65 connectivity studies include: 1) identifying genetically isolated populations of protected or endangered

66 species that should be monitored as separate conservation units (Palsbøll *et al.* 2007), 2) optimizing the size67 and location of protected areas to create well-connected reserve networks (Kritzer & Sale 2004; Jones *et al.*68 2007), 3) identifying stocks for fisheries management (Fogarty & Botsford 2007; Waples *et al.* 2008; Reiss *et*69 *al.* 2009), and 4) evaluating the potential impacts of resource exploitation on population dynamics of local70 communities (Bors *et al.* 2012; Plouviez *et al.* 2013). Besides these important applications to management

71 and conservation, connectivity studies can also help address long-standing questions. In ecology for

72 example they allow understanding to what extent marine populations are closed or open systems (Cowen *et*73 *al.* 2000), in microevolution they can be used to identify factors that create and maintain genetic74 differentiation (e.g. Bilton *et al.* 2002; Shank 2010), or to understand how local adaptation can occur in high75 gene flow species (e.g. Nielsen *et al.* 2009).

76 In this review, we underline the crucial importance for connectivity studies of the state of taxonomic

77 knowledge and correlated aspects on the available background knowledge on the biology of organisms.

78 These factors are sometimes overlooked despite the fact that they should condition the sampling design

79 and the inferences made from genetic data. One of the main considerations we develop is that properly

80 estimating connectivity at the population level requires assessing the robustness of available taxonomic

81 hypotheses. This is not trivial as for a large part of biodiversity, taxonomic knowledge is inadequate or even

82 lacking. Indeed, a great portion of the world’s biodiversity remains to be described, mainly in under-

83 explored, difficult-to-access habitats such as tropical rainforests and the marine environment (e.g.

84 Appeltans *et al.* 2012). In addition, in these habitats, species descriptions are often based on few specimens

85 and species distribution ranges and ecological requirements are poorly known. The development of DNA-

86 sequencing techniques has considerably accelerated the rate of discovery and the documentation of species

87 distributions. This is notably true in the marine realm, especially for poorly studied eukaryotic phyla for

88 which few or no other characters were available as reference for taxonomic delimitations (e.g. cryptic

89 lineages of macro- and micro-algae that are subject to convergent evolution towards reduced

90 morphologies; see for review Leliaert *et al.* 2014). However, turning these DNA-based discoveries (generally

91 only based on data such as those gathered from the Barcode of Life program), into robustly and formally  
92 named taxonomic entities remains a long and difficult process (Satler *et al.* 2013). Consequently, the  
93 literature includes a large number of undescribed and un-named cryptic or pseudo-cryptic species for which  
94 identification from one study to another is challenging in the absence of formal naming and description. As  
95 recently underlined by Fontaine *et al.* (2012), the time lapse between specimen collection and formal  
96 species description is extremely long (average “shelf life” of 21 years). A perhaps less appreciated issue is  
97 that even for abundant, commercially-important or ecologically well-studied taxa, nominal species were  
98 often described long ago and are not necessarily adequately delineated. This situation is particularly critical  
99 in the marine environment (e.g. Uthicke *et al.* 2010; Jaafar *et al.* 2012; Mantelatto *et al.* 2014; Thomas *et al.*  
100 2014). For instance *Ciona intestinalis* (Linnaeus, 1767), is a model organism in evolutionary developmental  
101 biology and phylogeny and was among the first animals to have its genome fully sequenced (Dehal *et al.*  
102 2002). However this taxon was only recently revealed as a complex of four cryptic species (Suzuki *et al.*  
103 2005; Caputi *et al.* 2007; Nydam & Harrison 2007; Zhan *et al.* 2010), including two widespread species that  
104 diverged ca. 3-4 my ago (Roux *et al.* 2013) but that are not yet formally named.

105 Beyond the discovery of recently diverged species, examples of taxonomic confusions at higher ranks are  
106 also not rare. For instance in 2009, Johnson *et al.* revealed that three families of deep-sea fishes artificially  
107 separated juveniles, males and females. This confusion was resolved by examining the morphology of more  
108 specimens and by using molecular data. This taxonomic revision, in which the three families are  
109 synonymized (the Cetomimidae being the valid family name), greatly changes the understanding of  
110 connectivity in terms of life history traits and ecology but should also affect the strategy for sampling  
111 populations.

112 The main objective of this review is thus to stress the fact that, when conducting connectivity assessments,  
113 the taxonomic status of the studied organisms should be critically reassessed, and revised if necessary, in  
114 light of newly acquired biological data. This point is particularly crucial in marine systems, which are the  
115 focus of this review, both because the marine biota accumulates the greatest number of deep evolutionary  
116 lineages (i.e. phyla), but also because it remains vastly under-sampled (see Costello *et al.* 2010 for  
117 multicellular organisms and Not *et al.* 2012 for the phytoplankton). Additionally, we outline that the type of  
118 data collected for estimating connectivity using population genetic methods may in turn be used to refine  
119 the taxonomic knowledge at the species rank. Actually, cryptic species are typically revealed by population  
120 genetic studies or by phylogeographic analyses. In fact, the patterns observed in connectivity studies and  
121 species delimitations belong to the same divergence continuum because both result from the same  
122 ecological and micro-evolutionary processes. Furthermore, the lack of knowledge for marine organisms is so  
123 important (e.g. Knowlton 2000 or Webb *et al.* 2010) that the question of species delimitations does not only  
124 concern the limits between recently separated lineages (i.e. the ‘emergent phylogenetic systems’ as defined

125 in Carstens *et al.* 2013) but also in the definition of higher-rank classifications as illustrated with the above  
126 example of Cetomimidae. We here first illustrate the importance of accurate taxonomic evaluation in the  
127 analysis of connectivity by examining one specific question (i.e. the correlation between pelagic duration  
128 and genetic structure estimates). We then highlight the main issues in relation with taxonomic assignment  
129 and propose a framework (guidelines) for preventing errors due to faulty taxonomy when conducting  
130 connectivity studies.

131

132 **An illustration of the problem: pelagic larval duration, genetic structure, and obscure species boundaries.**

133 The issue of poor species delimitations extends beyond individual studies when we consider meta-analyses  
134 or reviews investigating ecological trends such as the relationship between life-history traits and  
135 connectivity levels. A first, fundamental, step before conducting such comparative studies should be to  
136 ensure that comparisons are made across equivalent evolutionary units. Note that these evolutionary units  
137 may or may not be named. For instance, comparative studies including measures of genetic structure such  
138 as  $F_{ST}$  may produce biased conclusions if taxonomically-unresolved taxa are included. For instance in *C.*  
139 *intestinalis* the value of  $F_{ST}$  computed with microsatellites is six times higher between two populations, each  
140 made of the different cryptic species attributed to this name, compared to the values computed between  
141 populations including only one of these cryptic species (Zhan *et al.* 2010). If cryptic species are not treated  
142 as distinct entities, an analysis of genetic diversity among different sampling locations could lead to  
143 confound extrinsic barriers to dispersal with intrinsic reproductive barriers. We illustrate this issue by  
144 critically examining some of the source data used in a series of meta-analyses comparing measures of  
145 genetic differentiation across multiple species (Weersing & Toonen 2009, Selkoe & Toonen 2011). These  
146 meta-analyses address the question of whether early life-history traits influence connectivity and if this  
147 influence may be evidenced by the analysis of neutral genetic markers. This question is debated in marine  
148 molecular ecology because for most marine multi-cellular organisms, early life-history traits greatly differ  
149 from adult life-history traits (especially in terms of dispersal ability), and thus putatively affect connectivity  
150 among populations (especially if the habitat is fragmented). Different reviews and comparative studies have  
151 reached conflicting conclusions (e.g. Kinlan & Gaines 2003; Siegel *et al.* 2003; Shanks 2009; Weersing &  
152 Toonen 2009; Riginos *et al.* 2011; Selkoe & Toonen 2011). Weersing & Toonen (2009) examined how well  
153 pelagic larval duration (PLD) correlates with genetic estimates of connectivity. Their meta-analysis was  
154 based on a pseudo-random selection of papers from hits with selected keywords on electronic searches.  
155 The analyzed dataset included 130 species from 87 papers. Contrasting with the general expectation, the  
156 authors found only a weak correlation between PLD and genetic structure as estimated by the overall  $F_{ST}$ . As  
157 emphasized by Selkoe & Toonen (2011), this result may be explained by various factors, including the  
158 metrics used for estimating effective dispersal (i.e. gene flow), the type of markers and the sampling design.

159 In addition, we suggest that one possible explanation for such a weak correlation may be that unrecognized  
160 cryptic species within organisms with high PLD have been included in the dataset. Indeed, if two cryptic  
161 species are combined under a same name,  $F_{ST}$  estimates might be high not because the sampled  
162 populations are geographically structured (i.e. limited gene flow) but simply because these populations do  
163 not belong to the same species. At least some of the selected papers include questionable taxonomic  
164 entities. For example, the reef fishes attributed to five different *Halichoeres* Rüppell, 1835 species had  
165 among the highest  $F_{ST}$  values and unexpectedly the highest PLD. Examination of the available data from the  
166 literature for these species shows that different taxonomic treatments have been applied for similar  
167 divergence patterns in this genus (Rocha *et al.* 2005; Rocha *et al.* 2008). For example the name *Halichoeres*  
168 *bivittatus* (Bloch, 1791) covers two divergent mitochondrial lineages whereas the names *Halichoeres*  
169 *brasiliensis* (Bloch, 1791) and *Halichoeres radiatus* (Linnaeus, 1758) are attributed to two mitochondrial  
170 lineages separated by a level of divergence similar to that of the lineages within *H. bivittatus*. The two  
171 mitochondrial lineages of *H. bivittatus* are allopatric, except in Bermuda where both are present, although  
172 not in the same environments. The high overall  $F_{ST}$  within *H. bivittatus* might thus result from endogenous  
173 reproductive isolation barriers rather than dispersal barriers within a metapopulation. The case of the fish  
174 *Elacatinus evelynae* (Böhlke & Robins, 1968) is another example. Several mitochondrial and nuclear genetic  
175 markers, and an adequate sampling of closely related species revealed three distinct and well-supported  
176 phylogenetic lineages (Taylor & Helberg 2005) but the very high  $F_{ST}$  value used in the meta-analysis by  
177 Weersing & Toonen (2009) was estimated over a sample that combines these three lineages (Taylor &  
178 Heldberg 2003). Whether the three lineages evidenced by the phylogenetic approach need to be described  
179 as distinct species is an open question but an important criterion would be to determine if the divergence is  
180 definitive (i.e. definitive cessation of reproductive exchanges whatever the causes). In the case of  
181 reproductive isolation, the correlation between  $F_{ST}$  estimates and PLD should be evaluated only within each  
182 lineage. These examples illustrate that depending on how the available results are interpreted in terms of  
183 taxonomic status, conclusions of comparative studies on patterns of connectivity can vary widely.

184

### 185 **Theoretical framework for delimiting species: a primer.**

186 This critical analysis underlines a crucial aspect of taxonomy: described taxa are not facts but testable  
187 hypotheses about the structure of biodiversity. This is true not only for higher taxonomic ranks, generally  
188 considered as arbitrary ranks, but also for the species rank, which is the only taxonomic category for which a  
189 “biological reality” is recognized by most scientists (see for example in Mishler 2009). The question is then  
190 how such hypotheses are formulated in reference to a theoretical background and how they are then  
191 revised in light of new empirical data. The literature about the species concept is very large and the debate  
192 has often been considered as unsolvable (often referred as “the species problem”). In this debate, de



193 Queiroz (1998) was among the first to identify that two questions of different nature are mingled. He  
194 indeed underlined that the question of the definition of the species category has to be distinguished from  
195 that of the question of the adequacy of the criteria used in practice in the definition of the taxa ranked as  
196 species. The definition of the species category is an ontological question that may be addressed based on  
197 the theoretical ground offered by the Evolutionary Theory (as discussed in Samadi & Barberousse 2006).  
198 How each species-taxon should be delineated is an epistemological question that requires evaluating how  
199 well empirical data and analysis methods allow the taxonomists to propose sets of organisms that fit the  
200 chosen definition of the species category. Following this line of thinking, species-taxa are viewed as  
201 scientific hypotheses (species-hypotheses) engaged in a process of falsification based on the acquisition of  
202 new evidence. Although this analysis was only recently accepted among biologists, it closely corresponds to  
203 the working process of taxonomic revisions, which is one of the major activities of taxonomists (see  
204 discussion in Barberousse & Samadi, 2010). This analysis is also the conceptual background of the presently  
205 active field of integrative taxonomy. In the integrative approach to taxonomy, various attributes  
206 (morphology, DNA, geographical range, habitat, behavior, etc...) and criteria (i.e. the so-called "species  
207 concepts" such as the biological, the phylogenetic or the phenetic criteria), analyzed using various methods,  
208 are taken into account to propose robust species-hypotheses, reproducible and testable (see Camargo &  
209 Sites 2013 for a review on the integration of many characters and methods to delimit species). As  
210 underlined by Yeates *et al.* (2011), among others, the way hypotheses are validated should not be primarily  
211 the intersection among the evidence from different datasets but rather the prioritization of alternative  
212 species-hypotheses based on their explanatory power as defined within the framework of Evolutionary  
213 Theory. In this process the quality or robustness of available species-hypotheses greatly depends on the  
214 state of knowledge about the concerned compartment of the tree of life. For instance the state of art of  
215 species delimitations in birds is far more stable than that of most marine organisms. In fact uncertainties  
216 about available species-hypotheses may have two different causes. First, the state of art of taxonomy may  
217 be so preliminary that different life-stages of a same organism might have been erroneously classified in  
218 different species or even families (see above in Johnson *et al.* 2009). Second, the evidence that two  
219 evolutionary lineages are engaged in definitive divergence may not be easy to evaluate when the time  
220 elapsed since the split between the emergent lineages is recent (the "grey zone" of de Queiroz, 1998). If the  
221 relevance of taxonomic knowledge needs to be evaluated before conducting any biological study on a given  
222 organism we also should underline that the results of connectivity studies, if carefully analyzed, can in turn  
223 provide evidence about the state of divergence among emergent lineages.

224

225 **Where do we go from there?**

226 Four main sources of difficulty may strongly influence measures of connectivity and taxonomic hypotheses:  
227 (1) the state of knowledge about the biology of the studied organisms, (2) the state of taxonomic  
228 treatments of the studied organisms, (3) the spatial and temporal scales of sampling, and (4) the characters  
229 used to infer connectivity patterns. In the following section we review these four points and give some  
230 guidelines through examples taken from diverse marine model systems. We also offer some guidelines for  
231 designing connectivity studies when little background knowledge is available.

232  
233

## 234 **II. Methodological considerations and guidelines for connectivity studies.**

235

### 236 **1. Background knowledge on the biology of organisms.**

237 The availability of adequate background knowledge on the biology of organisms is crucial both for a robust  
238 taxonomic framework and for the study of connectivity. As discussed in the introduction, one of the life  
239 history traits commonly examined when studying connectivity data for marine organisms is pelagic larval  
240 duration (e.g. Cowen & Sponaugle 2009; Shanks 2009; Weersing & Toonen 2009; Selkoe & Toonen 2011).  
241 This biological character is also crucial for taxonomists when dealing with species delimitation (see an  
242 example in gastropods in box 5). However, other biological attributes such as reproductive strategies, sexual  
243 dimorphism or ecological specificities (i.e. habitat specialization, biological interactions, phenology, etc ...) can  
244 significantly influence both interpretations of connectivity data and taxonomic decisions. For example,  
245 advances in understanding of the migratory behavior of some focal marine organisms, such as eels, salmon  
246 or whales (e.g. Bottom *et al.* 2005; Aarestrup *et al.* 2009; Baker *et al.* 2013) revealed the existence of  
247 seasonal migrations among breeding and feeding sites, or discrete breeding sites, that have bearing on  
248 connectivity estimates. For instance, sex-specific philopatry affects the pattern of genetic structure in sperm  
249 whales (Engelhaupt *et al.* 2009) and humpback whales (Baker *et al.* 2013). In both cases, patterns of genetic  
250 differentiation are interpreted as a consequence of sex-biased philopatry, because it was possible to  
251 determine the gender of each individual analyzed, which is not possible for many poorly known organisms.  
252 The importance of sexual dimorphism, season and location of sampling can greatly affect the observed  
253 patterns of genetic differentiation and has consequence in the evaluation of species diversity (illustrated by  
254 the example of *Lessonia* in box 2). While the migratory behavior of some well-studied marine species has  
255 been characterized, the existence of seasonal migrations or discrete breeding sites is poorly known in most  
256 marine groups. Thus, the unreported occurrence of seasonal migrations may bias interpretations obtained  
257 when samples are collected from a single season. In extreme cases, different life history stages or sexes can  
258 wrongly be attributed to different taxa (e.g. Johnson *et al.* 2009), confounding estimates of connectivity.  
259 When the different sexes or stages inhabit different environments (such as different depths, e.g. the case of

260 the Cetomimidae in Johnson et al 2009), estimating dispersal from a single life-stage or sex can bias  
261 interpretation of connectivity patterns but also taxonomic interpretations. Other attributes, such as  
262 clonality or alternation of generations, can also complicate connectivity studies. In the case of clonal species  
263 such as seagrasses or some algae, neglecting such attributes would lead to a severe underestimation of  
264 gene flow or to overlook the presence of cryptic species (see in Box 4 and also Adjeroud *et al.* 2014 for  
265 another case in corals). Thus, the sampling design and interpretation of genetic data in connectivity studies  
266 and taxonomic revisions should take into account life history and ecological traits that could be relevant to  
267 estimating dispersal. If, as in many cases, such knowledge is limited or inexistent for the species of interest,  
268 one could examine information available for closely related taxa and/or incorporate this uncertainty in the  
269 sampling design (by sampling at different seasons, or different depths for example).

270

## 271 **2. Taxonomic knowledge.**

272 Connectivity studies generally consider the described species (i.e. available species hypotheses) as facts  
273 upon which the sampling scheme is established. In other words, the attributes associated with a species-  
274 hypothesis (either extracted from the original description or from other biological studies based on this  
275 species hypothesis) are considered as objective facts. Notably, the species range and habitat specificity are  
276 among the important attributes extracted from those hypotheses that are used in the design of a  
277 connectivity study. However, an important source of uncertainty in taxonomic hypotheses is the sampling of  
278 organisms and/or of characters upon which the species-hypotheses were initially drawn. Thus, data  
279 provided by a connectivity study are potentially a source of falsification of accepted taxonomic hypotheses.  
280 Consequently, for organisms for which no recent taxonomic revision is available, the sampling scheme  
281 should both follow the requirements of population biology but also that of phylogenetic systematics. For  
282 example, only a widespread sampling of a supposedly cosmopolitan species may reveal several diverging  
283 lineages restricted to isolated areas or distinct habitats as exemplified by the study of the supposedly  
284 ubiquitous and widespread amphipod *Eurythenes gryllus* (Lichtenstein in Mandt, 1822) (Havermans *et al.*  
285 2013) or the copepods *Pleuromamma piseki* Farran, 1929 and *Pleuromamma gracilis* Claus, 1863 (Halbert *et*  
286 *al.* 2013). Collecting in under-sampled or unexplored areas may also reveal new divergent lineages, like in  
287 the cosmopolitan red algae *Asparagopsis armata* Harvey, 1855 and *Asparagopsis taxiformis* (Delile) Trevisan  
288 de Saint-Léon, 1845, for which human-mediated transports increase the uncertainty about taxon history  
289 and systematics (Dijoux *et al.* 2014). Sampling is also an issue for species described as geographically  
290 restricted or ecologically specialized: looking for (and sampling) the targeted species or a priori closely  
291 related species in other localities or habitats may help rejecting hypotheses of endemism and/or  
292 specialization. For example, a deep-sea chemosynthetic mussel attributed to a new *Idas* Jeffreys, 1876  
293 species by Ritt *et al.* (2012), sampled in a deep cold seep site in the eastern Mediterranean sea, is included

294 in the same *cox1* cluster as the specimens attributed to *Idas simpsoni* (J. T. Marshall, 1900) and sampled at  
295 about 150 m depth in the north Atlantic on vertebrate bones (Thubaut *et al.* 2013). In this example the new  
296 species was considered as restricted to a given habitat that is highly fragmented and consequently other  
297 specimens sampled in a distant and distinct habitat were falsely assigned to a distinct species. Two sets of  
298 populations, identified as different species each associated with different habitats, may actually be  
299 populations of a single species connected through (even if rare) gene flow. Similar observations were also  
300 recently reported among chemosynthetic shrimps that were considered as distinct in cold seeps and vents,  
301 while molecular analysis revealed the occurrence of a single species with high gene flow across the Atlantic  
302 (Box 3; Teixeira *et al.* 2013). These examples illustrate the fact that the definition of the targeted group  
303 (ingroup) may be biased due to a lack of samples from specimens a priori identified as belonging to a  
304 different (related) species. Such sampling biases correspond to higher rank classification errors and may for  
305 example exclude from genetic analyses a subset of populations despite its crucial role in shaping  
306 connectivity patterns. As underlined below, the risk of inadequacy of available species-hypotheses should  
307 not be ignored and sampling cannot be limited to the intra-specific level. Because of the poor state of  
308 taxonomic knowledge for most marine organisms, taxonomic hypotheses (i.e. accepted classifications)  
309 should thus be carefully evaluated prior to assessing connectivity.

310

### 311 **3. Guidelines for spatial and temporal sampling of organisms.**

312 Sampling should be guided by two objectives: adequately covering (1) both the putative geographic and  
313 ecological range of the species of interest (coverage at the population level), and (2) potentially closely-  
314 related species over a wide geographic and ecological range (taxonomic coverage). Coverage at the  
315 population level should also take into account potential temporal variability in the geographic location of  
316 individuals (e.g. different location and time of breeding and feeding) and in habitats (e.g. pelagic *versus*  
317 benthic).

318 As explained above, this two-level sampling (developed in the case study presented in Box 5) should ensure  
319 a proper estimation of the upper bound of intra-specific diversity and the lower bound of inter-specific  
320 diversity as to determine their potential overlap. The genetic distance between conspecifics and individuals  
321 from different species, known as the “barcoding gap” (Meyer & Paulay 2005), is often used as a threshold to  
322 delimit species (at least in a primary step – Puillandre *et al.* 2012a) but its detection and meaning highly  
323 depend on the sampling scheme. These sampling guidelines are of course difficult to follow in all situations,  
324 especially in the marine realm: rarity, endemism, habitat fragmentation, human-mediated transports (i.e.  
325 biological introduction), lack of knowledge on sister-species relationships for most taxa, technical difficulties  
326 in sampling in the deep sea, among others factors, make the estimation of intra- and inter-specific diversity  
327 difficult. Such strategy may point to new species-hypotheses and thus to the need of taxonomic revisions.

328 However, such revisions (i.e. providing named species-hypotheses) are possible only if the requirements of  
329 the codes of biological nomenclature are taken into account in the sampling scheme. Indeed, an important  
330 requirement for attributing species names to the redefined species hypotheses is the inclusion of type-  
331 material linked to available species-names (see for example Puillandre *et al.* 2011). This last step is often  
332 lacking in molecular studies and the literature is spoiled by arbitrarily-labeled species hypotheses (e.g.  
333 OTUs, MOTUs, n. sp etc ...) that are difficult to compare from one study to another. As most name-bearing  
334 specimens (holotypes) cannot be integrated in molecular studies (due to their age, their poor conservation  
335 state, or simply because they were lost), sampling in type localities (as defined for example for animals in  
336 the article 76 of the International Code of Zoological Nomenclature) is a minimum requirement for naming  
337 species-hypotheses. This will both allow the corroboration (or the rejection) of available species-hypotheses  
338 and an adequate analysis of the pattern of connectivity within the targeted species. Implementation of the  
339 aforementioned sampling strategy and interpretation of the results should be guided as far as possible by  
340 this two-level method, especially when assessing the reliability of the inferred patterns of connectivity. In  
341 summary, these guidelines aim at identifying the adequate spatial and evolutionary scale for connectivity  
342 studies.

343

#### 344 **4. Guidelines for the sampling of characters.**

345 Classical markers widely used for marine organisms include genes from the phylogenetic literature and  
346 notably those mobilized in the DNA barcoding initiatives (typically, the *cox1* gene for metazoans, Bucklin *et*  
347 *al.* 2011; *cox1* or *RbcL* genes for macro-algae, Leliaert *et al.* 2014), but also other mitochondrial and nuclear  
348 genes such as ribosomal genes or “universal” introns (e.g. Jarman *et al.* 2002; Chenuil *et al.* 2010; Gérard *et*  
349 *al.* 2013). By using a set of unlinked genetic markers, it is possible both to refine the phylogenetic  
350 relationships (phylogenetic criterion) and to test whether specimens from the same species are able to  
351 recombine and, conversely, whether specimens from different species cannot (biological criterion). This can  
352 be achieved by sequencing, for each specimen, at least two unlinked markers, such as one from the  
353 mitochondrial genome (or chloroplast genome in plants and algae) and one from the nuclear genome. It is  
354 now widely accepted that the use of a single marker, especially mitochondrial, is risky when delimiting  
355 species, and should be used only in a first step to quickly propose primary species-hypotheses. As  
356 underlined in the review of Carstens *et al.* (2013), multi-locus approaches are required especially when  
357 considering recently diverging lineages, notably because of the genome wide heterogeneity of introgression  
358 rates observed between such genomes (e.g. between the two tunicates *C. intestinalis* type A and type B,  
359 Roux *et al.* 2013; or between the two sibling species of red algae *Gracilaria gracilis* (Stackhouse)  
360 M.Steentoft, L.M.Irvine & W.F.Farnham, 1995 and *Gracilaria dura* (C.Agardh) J.Agardh, 1842, Destombe *et*  
361 *al.* 2010). Furthermore, in some taxa (e.g. cnidarians, sponges, plants, algae...), classical markers sometimes

362 lack the necessary resolution to distinguish closely related species and to infer connectivity (e.g. Calderón *et*  
363 *al.* 2006; Shearer & Coffroth 2008). Examples can be found in metazoan groups with remarkably low levels  
364 of mitochondrial evolution such as corals, when this sole genetic marker is considered (e.g. Shearer &  
365 Coffroth 2008; McFadden *et al.* 2011). However, for many poorly studied organisms, identifying a nuclear  
366 marker variable at the species level in addition to, for example, the *cox1* gene can be tricky (e.g. McFadden  
367 *et al.* 2010). In such taxa, genetic similarity among individuals may reflect above-species evolutionary  
368 relationships rather than demographic connectivity (e.g. Miller *et al.* 2010). Clearly, these technical  
369 difficulties should be overcome thanks to the new methods based on Next Generation Sequencing, such as  
370 the RAD-sequencing (Baird *et al.* 2008). In a single run, many independent loci covering a large part of the  
371 genome are sequenced. They are thus particularly appropriate to test for recombination among individuals  
372 and to unravel diagnostic molecular characters among taxa (e.g. Pante *et al.* accepted).

373 Even if DNA characters often provide reliable data for species delimitation, either using a single locus to  
374 propose primary species-hypotheses (e.g. “species discovery” by DNA barcoding) or multiple loci to test for  
375 recombination among individuals within a species, they may not always provide decisive answers. When  
376 genetic data are not fully conclusive, morphological, geographical and ecological data can help turn the  
377 scales toward alternative hypotheses. For example Kekkonen & Hebert (2014) use a geographic criterion  
378 (*i.e.* genetic units are or not in sympatry) to conclude about the status of the detected genetic units. Our  
379 goal is not here to list all the characters and methods that can be included in the integrative taxonomy loop  
380 (as illustrated in Figure 1), but just to emphasize the need to discuss alternate species hypotheses, referred  
381 as (M)OTUs ((Molecular) Operational Taxonomy Units), ESU (Evolutionary Significant Units), PSH (Primary  
382 Species Hypotheses) or candidate species, depending on the method and characters, using all the available  
383 data to converge to robust, reproducible and testable hypotheses.

384

### 385 **5. Prioritization and workflow in multi-marker approaches.**

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

396 The principal source of difficulty for integrative taxonomy comes from cases of recent divergence  
397 ('emergent phylogenetic systems' as defined in Carstens *et al.* 2013). For example, under the hypothesis of  
398 a single species, high  $F_{ST}$  estimates are interpreted as population structure (low connectivity). If this species-  
399 hypothesis covers in fact two recently separated lineages, depending on the level of shared polymorphism  
400 among them, the  $F_{ST}$  estimates would be more or less close to one. Thus when taxonomy is uncertain, the  
401 two alternative hypotheses explaining such high  $F_{ST}$  estimates should be examined. Wang *et al.* (2008)  
402 illustrate this case by discussing how to discriminate between these two alternative hypotheses in their  
403 investigation of two morphotypes of finless porpoise (genus *Neophocaena*, Cetacea) that occur in sympatry  
404 in the strait of Taiwan. In their study, they apply population genetics methods to distinguish between a case  
405 of recent speciation and a case of low but recurrent gene flow. When simulating two populations with no  
406 gene flow during the last 18,000 years, simulated and observed  $F_{ST}$  and migration rate estimates were  
407 similar. Thus, the simulated scenario of a recent speciation event matched empirical data, leading the  
408 authors to conclude that the two morphotypes constitute two recent species that are currently in secondary  
409 contact. In this study the two species are not reciprocally monophyletic (phylogenetic criterion) but the  
410 population genetic approach supports the hypothesis of a complete breakage of gene flow (biological  
411 criterion) between the two morphotypes. Reciprocal monophyly is a criterion frequently used for species  
412 delimitation using molecular methods (e.g. Funk & Omland 2003) because it is much easier to evaluate than  
413 the biological criterion (inter-fecundity within species and inter-sterility among species). When the  
414 separation among studied species is ancient, the criterion of reciprocal monophyly is an operational method  
415 for detecting species. Conversely, among recently diverged species (i.e. 'emergent phylogenetic systems'),  
416 we recommend to give advantage to population genetic methods allowing the detection of lasting breakage  
417 of gene flow over the criterion of reciprocal monophyly. New methods of analysis, mostly based on  
418 coalescent theory, allow combining both criteria and should be as much as possible used in this context (e.g.  
419 Yang and Rannala 2010, Leaché *et al.* 2014, Hsieh *et al.* 2014). These methodologies coupled with new  
420 genotyping techniques allow the use, in a single analysis, of both the phylogenetic criterion (*i.e.* two sets of  
421 organisms have distinct evolutionary trajectories) and the biological criterion (*i.e.* the two sets of organisms  
422 are reproductively isolated) but were used only in a few pioneering studies (e.g. Leaché *et al.* 2014). Their  
423 applicability on poorly known organisms of the marine realm has yet to be explored but the robustness of  
424 such analyses will greatly depend on the quality of the sampling scheme relative to the state of taxonomic  
425 (and biological) knowledge.

426

### 427 **III Conclusion and implications.**

428

429 In many marine species, particularly those with high fecundity and large population sizes, population  
430 genetic studies have often revealed weak genetic structure, preventing accurate analyses of connectivity  
431 patterns. However, strong genetic structure patterns have also been revealed in marine systems. In this  
432 review, we argued that, in some cases, such observed patterns may actually be the outcome of incorrect  
433 sampling and/or incorrect taxonomic assignment.

434 We showed that unreliable definition of evolutionary units can lead to either overestimation or  
435 underestimation of connectivity. This situation is analogous to type I and type II statistical errors in a  
436 hypothesis-testing framework where full connectivity is the null hypothesis. In this review we illustrated two  
437 general cases, in which both types of errors may be encountered. These two cases correspond to the  
438 number of evolutionary units being (1) larger or (2) smaller than the number of species acknowledged in the  
439 study. In the first case, *i.e.* over-merging of evolutionary units, high genetic structure is equivalent to false  
440 positives and may emerge when populations from different, unrecognized or cryptic species are lumped  
441 together (e.g. Box 2 on the kelp genus *Lessonia* Bory de Saint-Vincent, 1825). This would lead to an increase  
442 in type I statistical error, in which the null hypothesis of a lack of barrier to gene flow among populations  
443 would be falsely rejected because species, and therefore populations, were not properly delimited. One  
444 simple example documented above is when falsely inflated  $F_{ST}$ -values due to mixing of specimens from  
445 different evolutionary entities are interpreted as picturing limited dispersal and increased genetic drift. The  
446 issue lies in the interpretation of the data rather than in the data themselves, and some examples were  
447 discussed above (e.g. Rocha *et al.* 2005; Taylor & Helberg 2005; Rocha *et al.* 2008). The second case, *i.e.*  
448 over-splitting of evolutionary units, results in the flawed *a priori* assumption of a complete lack of  
449 connectivity among taxa described as distinct species (see the example of the Atlantic shrimp named,  
450 depending on its habitat, *Alvinocaris muricola* Williams, 1988 or *Alvinocaris markensis* Williams, 1988 as  
451 detailed in Box 3), preventing their recognition as potentially interdependent entities and the will to engage  
452 in connectivity studies among their populations.

453

454 As illustrated by the examples discussed in this paper, the issue of unreliable delineations of evolutionary  
455 units has been exacerbated in the past two decades with the rise of molecular data, particularly in the  
456 framework of phylogeographic studies documenting an ever-increasing number of cryptic species in the  
457 marine realm (Appeltans *et al.* 2012). Species are hypotheses that need to be revised if necessary in light of  
458 new information. To anticipate the possibility of false taxonomic hypotheses, we highlighted several  
459 important aspects. First, the sampling design should encompass as much as possible 1) the entire known  
460 distribution range of the targeted organism and 2) closely related taxa, or even some distantly related taxa  
461 in taxonomic groups where the phylogeny is poorly known (e.g. deep-sea organisms). The latest point is  
462 rarely considered in connectivity studies although it is an efficient way to ensure the correct definition of



463 the focal (i.e. ingroup) species. Second, any connectivity study should be careful in its design regarding life-  
464 history traits: pelagic larval duration, reproductive system characteristics (sex-ratio, clonal reproduction,  
465 ...etc) and social organization, which all have important consequences on connectivity patterns but also on  
466 species delimitations. Third, the types of markers used should be diversified; in particular the sole use of  
467 mitochondrial genes should be avoided. Altogether, the three points mentioned above are all relevant to  
468 the framework of integrative taxonomy. This approach has proven to be helpful in species delineation and  
469 meaningful when coupled with or used for connectivity studies.

470 Importantly, connectivity studies can reciprocally provide new data that can help defining new species in  
471 previously recognized nominal species. This is particularly well illustrated in the case of introduced or  
472 invasive marine species (see box 6). Human-driven dispersal adds to natural dispersal, thereby broadening  
473 species ranges, sometimes at a worldwide scale. Connectivity and phylogeography studies of marine  
474 invaders have shown that important and cosmopolitan invaders are actually species complexes (e.g. in  
475 colonial tunicates *Botryllus schlosseri* (Pallas, 1766), Bock *et al.* 2012; in solitary tunicates *C. intestinalis* Zhan  
476 *et al.* 2010; in the red alga *A. taxiformis* Dijoux *et al.* 2014). Cryptogenic, pseudo-indigenous species and  
477 cryptic species, as defined by Carlton (2009) are new issues deserving integrative taxonomy and barcoding  
478 studies, not only to verify the accuracy of invaders reports (e.g. McGlashan *et al.* 2008) or to quickly report  
479 newly introduced species (Bishop *et al.* 2013), but also to examine gene flow between cryptic species not  
480 fully reproductively isolated, in particular when invaders are morphologically similar to native species (e.g.  
481 Nydam & Harrison 2011). More generally, delineating species can be particularly challenging considering  
482 how rapidly species distribution can shift due to either introduction processes or global change. These  
483 issues increase the complexity of taxonomic assignment and biogeography studies.

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

492

#### 493 **Acknowledgments.**

494 The authors thank the participants and organizers of the French Marine Connectivity (GDR MarCo) working  
495 group, who stimulated discussions that were the basis of this review. GDR MarCo was funded by the Institut  
496 Français de Recherche pour l'Exploitation de la Mer (IFREMER) and the French Centre National de la

497 Recherche Scientifique (CNRS). Salaries for E.P. and A.V. were covered by a grant to the Poitou-Charentes  
498 region (Contrat de Projet-État-Région 2007–2013), by a grant from the Fond Européen de Développement  
499 Régional (EP), and by Actions Internationales et Rayonnement of the University of La Rochelle (AV). This  
500 review benefited from the deep-sea studies developed in the French-Taiwanese project TF-DeepEvo  
501 funded by ANR and NSC (ANR 12-ISV7-0005-01) and from the project CONOTAX, funded by the French ANR  
502 (ANR-13-JSV7-0013-01).

For Review Only

503 **Boxes.**

504 **Box 1. Species delineation blurred by hybridization and suboptimal sampling design: the case of *Fucus***  
505 **Linnaeus, 1753.**

506

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

512 The two closely related brown algae morphological species: *Fucus spiralis* Linnaeus, 1753 and *Fucus*  
513 *vesiculosus* Linnaeus, 1753 are found in sympatry along eastern Atlantic rocky shores, from northern  
514 Norway to Morocco. They differ mainly by their morphology, their vertical distribution on the shore and  
515 their reproductive system. *F. spiralis* is hermaphrodite and is found in the higher intertidal, whereas *F.*  
516 *vesiculosus* is dioecious and is located just lower on the shore. Molecular analyses of internal transcribed  
517 spacer of nuclear ribosomal DNA (ITS) were not efficient to separate these two taxa (Serrão *et al.* 1999) and  
518 their taxonomic status as two separate species was questioned. The use of five microsatellite loci on six  
519 populations of *F. spiralis* and ten of *F. vesiculosus*, separated by tens to hundreds of kilometers, showed that  
520 within each species individuals share diagnostic alleles whatever their geographic origin (Billard *et al.* 2005).  
521 Consequently, the phylogeographic approach of Billard *et al.* (2005) demonstrated that these two  
522 morphological species correspond to two different genetic entities. Using gene admixture detection  
523 methods, intermediate individuals between the two species were detected along shores where the  
524 distribution of the two species overlaps at the scale of few hundred meters (Engel *et al.* 2005). The authors  
525 suggested a recent divergence of the two taxa with retention of ancestral polymorphism or introgressive  
526 hybridization with varying levels of admixture.

527 Interestingly, using a more refined sampling (combining transects across the zonation of the shore and  
528 quadrats at the different tidal levels), a new genetic entity previously overlooked by Engel *et al.* (2005) was  
529 detected within the *F. spiralis* morphotype (Billard *et al.* 2010). These two entities were first labeled *F.*  
530 *spiralis* Low and *F. spiralis* High according to their vertical distribution on the shore. Based on an integrative  
531 approach combining molecular, physiological and morphological analyses, Zardi *et al.* (2011) proposed to  
532 elevate the new genetic entity, *F. spiralis* Low, to the species level using the name *F. guiryi* G.I.Zardi,  
533 K.R.Nicastro, E.S.Serrão & G.A.Pearson, 2011. The range distribution of *F. guiryi* was shown to occur from  
534 Brittany to Morocco (Coyer *et al.* 2011). Common garden experiments performed on the three sympatric  
535 genetic taxa: *F. spiralis*, *F. guiryi* and *F. vesiculosus* showed that physiological response to desiccation stress  
536 differed between species and was consistent with their respective vertical distribution on steep

537 environmental clines in exposure time (Zardi *et al.* 2011). The opportunity of hybridization between the  
538 three species was confirmed to be higher when they are physically close to each other on the shore while in  
539 southern Portugal, where *F. vesiculosus* is only found in estuaries and *F. guiryi* on the open coast, the two  
540 species never hybridize (Zardi *et al.* 2011). In addition to their ecology, it was proposed that selfing as well  
541 as limited gamete dispersal contribute to species integrity in the hermaphroditic species *F. spiralis* and *F.*  
542 *guiryi* (Engel *et al.* 2005; Billard *et al.* 2010; Zardi *et al.* 2011). Finally, a high level of asymmetric  
543 introgression was reported between these different species using a combination of nuclear and  
544 mitochondrial DNA markers (Coyer *et al.* 2011). A scenario of organelle capture via hybridization and spread  
545 of neutral nuclear alleles during range expansion was suggested to explain the common occurrence of  
546 introgression in this genus. This example clearly showed that organellar loci as barcoding tools should be  
547 used with care because of introgression.

548 **Box 2. The integrative taxonomy loop: study case of two cryptic species in the kelp *Lessonia nigrescens***  
549 **Bory de Saint-Vincent, 1826 complex.**

550 The discovery of cryptic species does not always lead to a taxonomic reexamination and in many cases,  
551 cryptic species remain without taxonomic characterization (and therefore with no corresponding Latin  
552 name) long after their detection. Here we describe the case of the kelp *Lessonia nigrescens* Bory de Saint-  
553 Vincent, 1826 complex that was recently identified as two separate cryptic species thanks to a considerable  
554 sampling effort, combining phylogenetic, population genetics and ecological approaches. This kelp forms  
555 forests along the intertidal South Eastern Pacific coasts and was long considered as a unique species from  
556 Cape Horn in southern Chile (ca. 55° S) to central Peru (ca. 12° S; Santelices 1989). Two cryptic species were  
557 recently identified through a multigene phylogeographic study (using a combination of four markers located  
558 in the three genomic compartments - chloroplast, mitochondrion and nucleus - with 1,000 individuals  
559 covering more than 2,500 km of coastline) (Tellier *et al.* 2009). The two divergent genetic lineages show a  
560 parapatric latitudinal distribution: one species extends from southern Peru (17°S) to central Chile (30° S),  
561 and the other extends from central Chile (29° S) to Chiloe Island (42°S). Between 29° S and 30° S there is a  
562 narrow area where both lineages spatially overlap in discrete patches where individuals belong to either the  
563 northern or southern species. This contact zone between 29° S and 30° S has been described as a mosaic of  
564 sites occupied by one species or the other (Tellier *et al.* 2009), rather than a true gradual transition zone,  
565 with one species replacing the other. In addition, a detailed population genetic study of gene flow in the  
566 transition zone confirmed a total absence of hybrids and mixed populations (Tellier *et al.* 2011). Ecological  
567 differentiation between the two cryptic species was investigated, using different approaches (e.g. controlled  
568 response to different thermal and desiccation stresses in the lab, Oppliger *et al.* 2012; López-Cristoffanini *et*  
569 *al.* 2013). Both microscopic (haploid) and macroscopic (diploid) phases of the haploid-diploid kelp life cycle  
570 were studied. Results demonstrated adaptive divergence for both phases that could explain the geographic  
571 segregation of these two cryptic species. In addition, contrasted demographic processes could be observed  
572 at various stages of their life cycle: female gametophytes from the Northern species develop rapidly but  
573 give very few oogonia per gametophyte, whereas female gametophytes from the Southern species delay  
574 their maturation, grow vegetatively and then produce numerous oogonia per individual. These results  
575 strongly advocate that, even though the origin of the genetic differentiation of the two species is ancient,  
576 ecological and intrinsic differences are probably still important for the maintenance of reproductive  
577 isolation.

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

581

582 **Box 3: Taxonomic inflation results in the flawed *a priori* assumption of a complete lack of connectivity**  
583 **among morphologically distinct populations: an example with shrimps in the deep-sea environment.**

584 As reported in this manuscript, most examples of taxonomic bias leading to distorted connectivity  
585 inferences are linked to the existence of hidden cryptic species or lineages. We here illustrate the opposite  
586 case in which the polymorphism of morphological characters led to taxonomic inflation and thus to a false  
587 inference of the connectivity pattern. This situation was found in the Alvinocaridae, a family of caridid  
588 shrimps specialized in deep chemiosynthetic ecosystems. In the Atlantic, two species were described in  
589 1988 by Williams: *A. muricola* associated with cold seeps from Western Atlantic and *A. markensis* from  
590 hydrothermal vents of the Mid Atlantic Ridge. The two species were distinguished based on a slight  
591 difference in an abdominal segment being entire or obscurely serrate. This distinction was made on juvenile  
592 specimens and later confirmed on adults (Komai & Segonzac 2005), although on the basis of a distinct  
593 morphological criterion that was the more inflated and convex nature of anterior part of the branchial  
594 region in *A. muricola* compared to *A. markensis*. In 2005, Komai & Segonzac cautiously identified as *A. aff.*  
595 *muricola* shrimps sampled along the continental margins of Eastern Africa. These authors extensively  
596 discussed the pan Atlantic distribution of *A. muricola* given the puzzling absence of geographically  
597 intermediate population. These authors hypothesized either the occurrence of large-scale dispersal  
598 capacities or the existence of unknown suitable sites in the intermediate area. A barcoding approach based  
599 on the Cytochrome Oxidase I gene (*cox1*), initially aiming at retracing the phylogenetic history of the  
600 alvinocarid shrimp, casted doubt on the taxonomic status of those two species, showing entire sequence  
601 match among haplotypes of both species (Teixeira *et al.* 2013). Indeed, using both mitochondrial and  
602 nuclear sequences (*cox1* and 18S) together with a set of microsatellite loci, Teixeira *et al.* (2013) showed  
603 that all populations respectively attributed to these two names share a unique set of alleles at both types of  
604 genetic markers and display extremely weak genetic differentiation. This result eventually confirmed the  
605 hypothesis of intermediate populations raised by Komai & Segonzac (2005) while more importantly  
606 demonstrating that *A. markensis* and *A. muricola* should not be considered as two distinct species.  
607 Following this interpretation the two species have to be synonymized in a single species that has a wide  
608 geographic distribution but also a diversified range of suitable habitats in the deep sea. This example  
609 supports the recommendation of investigating the phylogenetic framework of taxa studied including  
610 closely-related taxa before engaging into comprehensive connectivity studies.  
611 The same barcoding analysis also revealed that classification at both genus and species level should be  
612 revised within the Alvinocaridae with possible cases of synonymy of species described as distinct genera  
613 (e.g. *Rimicaris hybissae* Nye, Copley & Plouviez, 2012 reported in Eastern Atlantic seeps and *Chorocaris*  
614 *chacei* (Williams & Rona, 1986) and Mid-Atlantic vents).

615 **Box 4. The implications of working on organisms with clonal populations and alternation of generations:**  
616 **the case of seagrass meadows and algae.**

617

618 Seagrass meadows were initially expected to exhibit large-scale panmixia due to the positive buoyancy of  
619 their seeds and rhizomatic fragments. In studying connectivity in those partially clonal organism, the first  
620 issue was the ability to recognize clonal lineages on the basis of their multi locus genotypes (MLG) in order  
621 to discard from the datasets the sampling units that are replicates of the same clonal lineage that would  
622 potentially artificially inflate structure estimates such as  $F_{ST}$  among meadows (Arnaud-Haond *et al.* 2005;  
623 Arnaud-Haond *et al.* 2007a). Even avoiding this potential bias, however, population genetic analysis across  
624 the range area of several seagrass populations suggested a systematic structure and limitation to gene flow  
625 across all geographical scales among almost all discrete meadows analyzed (Coyer 2004; Olsen, 2004,  
626 Arnaud-Haond *et al.* 2007b; Alberto *et al.* 2008). Secondly however, fine grained analysis of the  
627 distribution of clonal lineages and genetic polymorphism in *Zostera marina* Linnaeus, 1753 meadows  
628 revealed a large influence of clonality, leading to a mosaic pattern of genetic patchiness largely driven by  
629 patchy events of recolonization followed by the occupation of space through clonal spread, despite the  
630 likely occurrence of much larger scale effective migration (Becheler *et al.* 2010; Becheler *et al.* 2014). These  
631 results are casting doubts on the accuracy of the genetic concept of populations in those partially clonal  
632 organisms, suggesting the ecological concept of distribution continuity (therefore the meadow scale) may  
633 be more accurate as a framework for connectivity analysis (Becheler *et al.* 2010) and genetic differentiation  
634 may not systematically reflect lack of connectivity. Additionally, the general lack of isolation by distance  
635 patterns reported across seagrass species was interpreted as a clue for scarce but significant events of large-  
636 scale dispersal influencing the dynamics, resilience and geographic distribution of genetic diversity in  
637 seagrasses (Kendrick *et al.* 2012). Whereas this may hold for short-lived species with a regular turn over and  
638 propagule recruitment, the decomposition of the genetic distance spectrum (Rozenfeld *et al.* 2007) among  
639 meadows of the long-lived seagrass *Posidonia oceanica* (Linnaeus) Delile, 1813 (Arnaud-Haond *et al.* 2012)  
640 also brings additional warning for the interpretation of genetic structure patterns in terms of contemporary  
641 gene flow. This approach of genetic data performed at the level of clonal lineages indeed showed that at the  
642 range scale, the pattern of genetic structure was reflecting ancient events of (re)colonization followed by  
643 long term clonal spread and accumulation of divergent mutations within meadows at the regional scale,  
644 rather than contemporary patterns of connectivity (Arnaud-Haond *et al.* 2014). Interestingly, in haploid-  
645 diploid species like many seaweeds, there is an alternation of haploid gametophytes and diploid  
646 sporophytes to complete the sexual life cycle. In such life cycle, dominance of clonal propagation can led to  
647 an uncoupling of haploid and diploid phases and thus to differences in allele frequency between  
648 gametophytes and tetrasporophytes (Sosa *et al.* 1998). This has been observed in species of red alga

649 reproducing mainly asexually (in *Gelidium arbusculum* Bory de Saint-Vincent ex Børgesen, 1927; Sosa *et al.*  
650 1998; *Gracilaria chilensis* C.J.Bird, McLachlan & E.C.Oliveira, 1986; Guillemin *et al.* 2008) while the  
651 occurrence of sex, even occasionally, in other species, explained the lack of significant differences between  
652 haploid and diploid subpopulations (*Gelidium canariensis* (Grunow) Seoane Camba in Sosa & Garcia-Reina  
653 1993; *Gracilaria gracilis* (Stackhouse) M.Steentoft, L.M.Irvine & W.F.Farnham 1995 in Engel *et al.* 2004;  
654 *Cladophoropsis membranacea* (Hofman Bang ex C.Agardh) Børgesen, 1905 in van der Strate *et al.* 2002 and  
655 *Chondrus crispus* Stackhouse, 1797 in Krueger *et al.* 2011). Genotyping both haploid and diploid individuals  
656 in these species is necessary in order to detect the clonality and its effect on population structure.

For Review Only



657 **Box 5. Large taxonomic and geographic sampling in the deep-sea gastropod *Bursa* Röding, 1798 helps**  
658 **delineating species before testing hypotheses on modes of speciation.**

659 Deep-sea seamounts offer a multi-scale fragmented habitat for benthic organisms. For gastropods  
660 inhabiting these environments, connectivity patterns have been shown to widely depend on the type of  
661 larval development (Castelin *et al.* 2010; Castelin *et al.* 2012). Notably, the three deep-sea Bursidae species  
662 (*Bursa latitudo* Garrard, 1961, *Bursa fijiensis* (Watson, 1881) and *Bursa quirihorai* Beu, 1987) share a largely  
663 overlapping geographic range on the seamounts of the Norfolk and the Lord Howe ridges, and the  
664 continental slopes of New Caledonia, in the Southwest Pacific, and are not genetically structured over this  
665 area.

666 The prevalence of allopatric versus sympatric modes of speciation, were studied following Coyne & Orr  
667 (2004)'s biogeographic approach that involves studying current distribution patterns and phylogenetics of  
668 recently-separated sister species. In this approach, one of the first requirements is to show that the species  
669 pairs have a sister relationship and that reproductive isolation is complete (i.e. absence of hybridization  
670 between the sister-species). All Bursinae species have a long-lived planktonic-feeding larva and most of  
671 them have a wide distribution range. Moreover, most are geographically rare and the distribution of  
672 populations within the species range is usually sparse and highly fragmented. This leads to confusions in  
673 Bursidae taxonomy with some geographically-isolated populations attributed to distinct species names, or  
674 slightly distinct morphotypes collected at distant but not geographically isolated sites attributed to the  
675 same species name. The sampling strategy was thus designed to meet a three-fold objective: extending the  
676 taxonomic sampling without any *a priori* hypotheses about the systematics of the group, at both the genus  
677 and the species level (e.g. including all available individuals, populations, species and genera of the family  
678 Bursidae); extending the geographic sampling toward other oceans (Atlantic, Indian); and extending the  
679 ecological sampling by adding species from both shallow-water and deep-sea environments.

680 Starting from this large taxon sampling, "Primary Species Hypotheses" (PSH) were drawn from the shell  
681 morphology using the taxonomic literature and dry material of the Malacology collection at the French  
682 National Museum of Natural History (MNHN). ESUs were then defined using both the GMYC (Pons *et al.*  
683 2006) and ABGD methods (Puillandre *et al.* 2012b) applied on *cox1* sequences and then strengthened with  
684 the analysis of a second mitochondrial gene (16S rRNA) and the nuclear 28S rRNA gene fragment. All  
685 sources of evidence generated from these four steps were integrated to propose "Retained Secondary  
686 Species Hypotheses" (RSSHs).

687 Over the sampled geographical range, the three morphologically diagnosable entities, *B. latitudo*, *B.*  
688 *quirihorai* and *B. fijiensis*, are reciprocally monophyletic based on all tested loci. The dataset also revealed  
689 cryptic lineages emphasizing that species diversity in the family Bursidae is underestimated and suggesting  
690 that a taxonomic revision of the family should be performed. Among New Caledonian species, *B. fijiensis*

691 and *B. quirihorai* are reciprocally monophyletic. These two species are the two most closely related species  
692 in the inferred phylogeny. The current biogeographic ranges of these two species and the estimated time of  
693 divergence make the scenario of sympatric speciation the most likely.

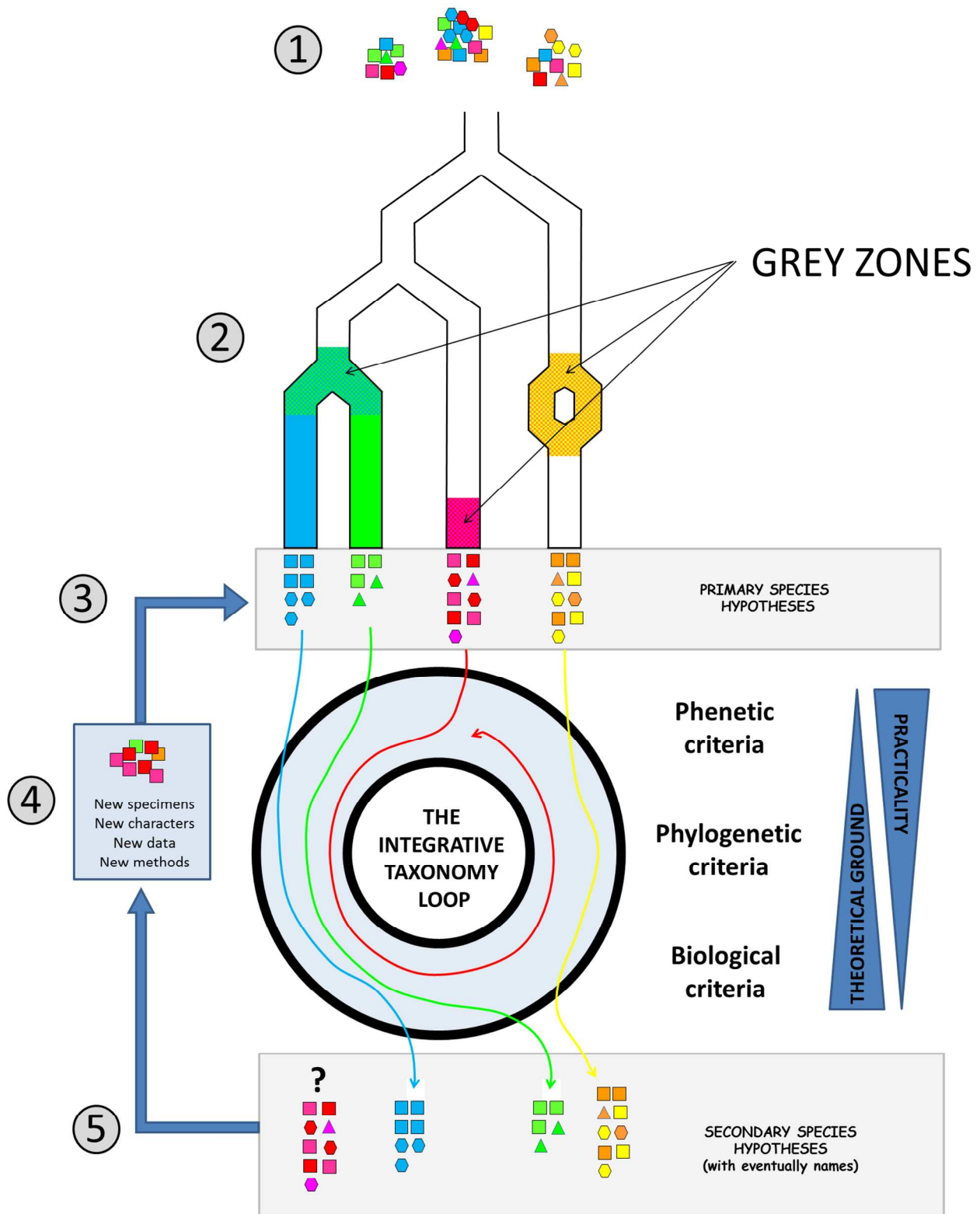
For Review Only

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

696 Besides the natural processes of expansion and contraction of the range colonized by a species over  
697 evolutionary times, sudden species-range expansions -often through human-mediated transport- have been  
698 observed at an increasing rate since the end of the 19th century. Since the original work by Elton published  
699 in 1958 (Elton 1958), biological invasions have received a growing set of attention from the scientific  
700 community (e.g. Blackburn *et al.* 2011 and Simberloff *et al.* 2013 for recent reviews and references herein).  
701 Tightly linked to human activities particularly through shipping and aquaculture, biological invasions  
702 severely alter biodiversity of marine ecosystems (Rilov & Crooks 2009). One of the important outcomes of  
703 marine biological invasion studies has been to document biotic homogenization at a global scale with the  
704 establishment of 'cosmopolitan' species and cryptogenic species (i.e. species that are neither demonstrably  
705 native nor demonstrably introduced as defined by Carlton (1996; Haydar 2012). Interestingly, many of these  
706 cosmopolitan invasive species actually appear to be composed of morphologically close taxa forming a  
707 species complex. This was recently exemplified in two model ascidians, the colonial tunicate *Botryllus*  
708 *schlosseri* (Bock *et al.* 2012) and the vase tunicate *Ciona intestinalis* (Nydam *et al.* 2011; Zhan *et al.* 2010).  
709 Both taxa are distributed worldwide. Phylogeography, phylogeny, population genetics and genomic studies  
710 showed for both taxa that they were actually characterized by marked genetic subdivision and large  
711 genealogical divergence (up to 4-8 million years for type A and type B in *C. intestinalis*, Roux *et al.* 2013).  
712 These two taxa are composed of (partially) reproductively isolated species, some of them came recently  
713 into secondary contact because of their introduction with human activities. This is notably the case for two  
714 members of the *Ciona intestinalis* species complex called type A and type B, which are living in sympatry  
715 after the presumably recent introduction of Type A in the Northern Atlantic; the two species are  
716 incompletely reproductively isolated in the English Channel where they are hybridizing (Nydam & Harrison  
717 2011). The outcome of such secondary contact is unknown but may promote either adaptive introgression  
718 or speciation (Abbott *et al.* 2013). Bock *et al.* (2012) also pointed out that the different cryptic species of the  
719 *Botryllus schlosseri* complex are not characterized by similar invasive abilities questioning the reasons for  
720 such differential abilities. Integrative taxonomy and molecular approaches are thus critical research to carry  
721 out in biological invasions. They may help identify new cryptic introduced species with different invasion  
722 abilities, understand the evolutionary history of cryptogenic species (e.g. the tunicate *Molgula manhattensis*  
723 *manhattensis* (De Kay, 1843) (Haydar *et al.* 2011) and the fate of secondary contact between previously  
724 isolated taxa. For management purpose, they are also helpful to identify new invaders (Bishop *et al.* 2013)  
725 and verify the accuracy of published records of invasive species (e.g. McGlashan *et al.* 2008) - especially for  
726 morphologically cryptic cosmopolitan taxa.

727 **Figure1: the integrative taxonomy loop**

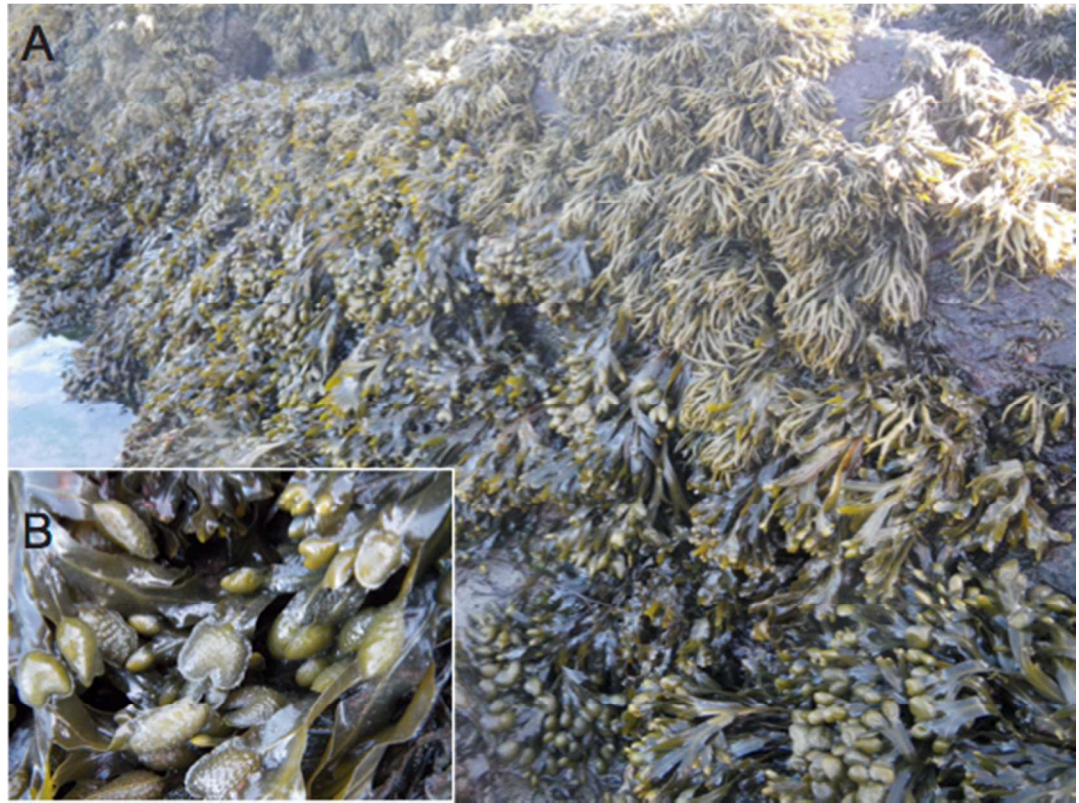
728



729

730

731 **Box figure 1.**  
732



733  
734

View Only

735 **Box figure 2.**

736

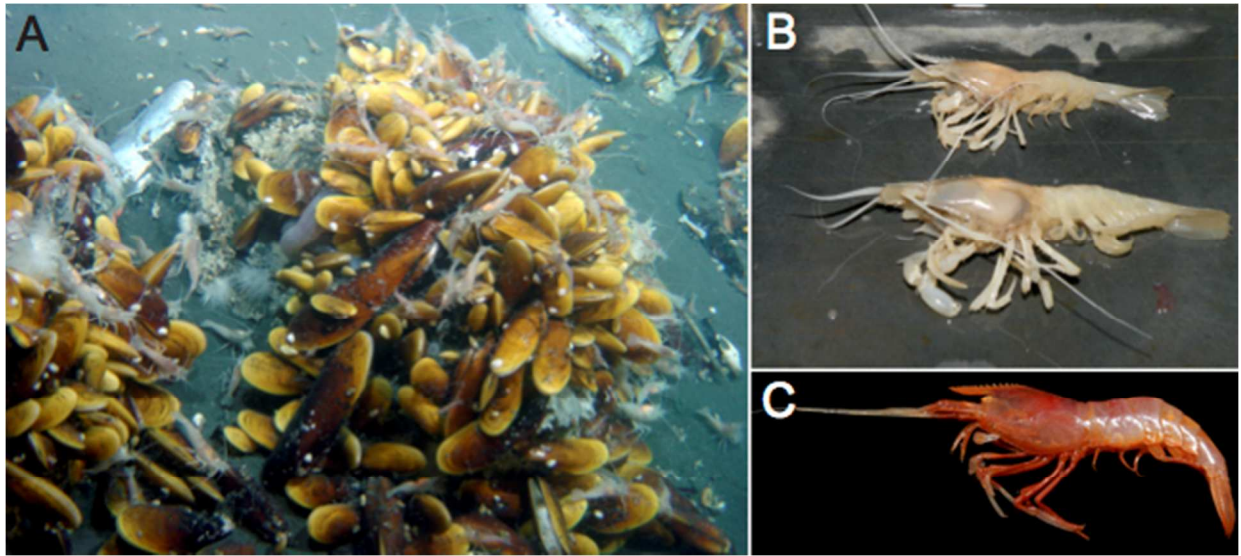


737

738

ew Only

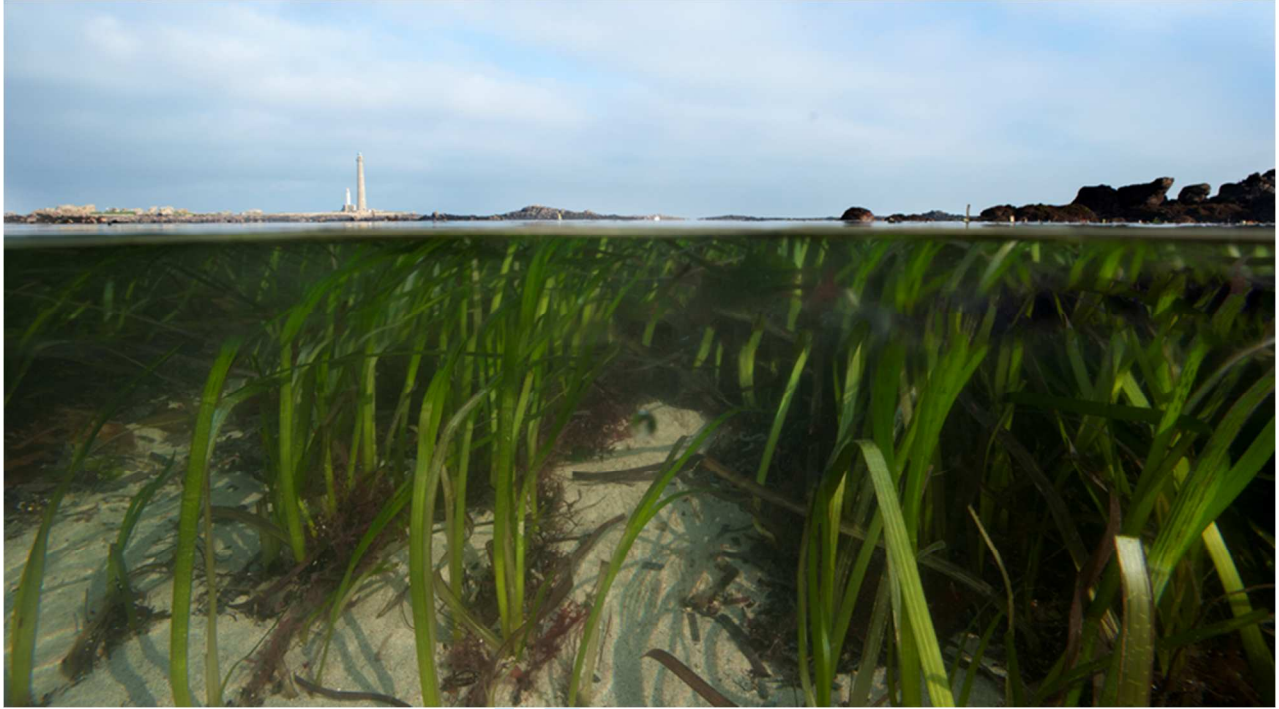
739 **Box figure 3.**  
740



741  
742

For Review Only

743 **Box figure 4.**  
744



745  
746

Review Only



747 **Box figure 5.**  
748



749  
750

Review Only

751 **Box figure 6.**  
752



©Wilfried Thomas – Station Biologique Roscoff

753  
754

755 **Figure captions**

756

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

777

778 **Box figure 1. Species delineation in *Fucus*.** A: Vertical distribution of furoid species on rocky shore from  
779 high to low-water mark. Site: Viana do Castelo (GPS coordinates: 41° 41' 51''N, 8° 51' 07'' W, northern  
780 Portugal), photograph: Christophe Destombe. B: Detail of reproductive structure of *Fucus guiryi* showing a  
781 typical receptacle sterile rim (arrow). Photograph: Christophe Destombe.

782

783 **Box figure 2. The pseudocryptic kelp species *Lessonia spicata*.** A: This species (on the left of the picture) is  
784 found in central-south Chile (29°-41°S) in areas exposed to wave actions. Sporophyte individuals can be  
785 more than 2 meters long and are anchored to the rock by a holdfast. Site Los Molles (GPS coordinates:  
786 32°14'S-71°31'W, central Chile), photograph: Pablo Balzo. B: The gametophytes of this species are  
787 microscopic and show a sexual dimorphism with males being smaller and more branched than females.  
788 Photograph: Valeria Oppliger.

789

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

795

796 **Box figure 4. *Zostera marina* meadows in Brittany.** Individual shoots (ramets) are easily distinguished while  
797 rhizomatic connections linking ramets belonging to the same clonal lineage (genet) are hidden below the  
798 sandy sédiment (photograph: Ifremer/Olivier Dugornay).

799

800 **Box figure 5. The deep-sea gastropod *Bursa*.** A specimen of *Bursa sp.* sampled in New Caledonia using a  
801 dredge. The other organisms sampled in the same dredge are typical of hard bottoms (e.g. stylasterids,  
802 deep-water solitary stony coral, etc...). Photograph MNHN/Pierre Lozouet.

803

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

806

807

808

809 **Glossary.**

810 **Biological nomenclature:** set of rules that determine how names are given to taxa. In the binominal system,  
811 when a new species is described based on a given set of biological criteria applied over a set of  
812 specimens, one of them is designated as the holotype (the name-bearing specimen). When the  
813 taxonomy is revised using new material (additional specimens) and/or new biological criteria the  
814 position of the name-bearing specimens corresponding to the available names for the studied taxa  
815 is used to name the new taxonomic hypotheses. When it is not possible to study the holotype for  
816 the new character being proposed (e.g. molecular data), other categories of type material might be  
817 used (paratypes, syntypes, etc...)

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

821 **DNA barcoding:** identification of an unknown specimen by comparison of a standard DNA fragment (e.g.  
822 Barcode fragment of the *cox1* gene for metazoans) to a database of sequences linked to identified  
823 vouchers. By extension, use of a standard DNA fragment for biodiversity studies (e.g. to discover  
824 new species; see Collins & Cruickshank (2013) for a distinction between “specimen identification”  
825 and “species discovery and delimitation”).

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

829 **Evolutionary significant unit (ESU):** defined in 1994 by Moritz to explicitly take into account the genetic  
830 component of biodiversity in conservation, ESU is defining historically isolated groups of  
831 populations showing reciprocal monophyly and significant genetic divergence at nuclear loci.

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

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

839 **Phylogeography:** the spatial distribution of phylogenetic lineages, particularly within and among closely  
840 related species (Avice 2009).

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

843 **Pseudocryptic species:** morphologically overlooked species distinguished using other lines of evidence (e.g.  
844 DNA) that proved to be finally distinguished by subtle morphological differences (e.g. Box 2 on the  
845 kelp *Lessonia*).

For Review Only

846 **References**

847

848 Aarestrup K, Økland F, Hansen MM *et al.* (2009) Oceanic spawning migration of the European eel (*Anguilla*  
849 *anguilla*). *Science*, **325**, 1660.

850 Abbott R, Albach D, Ansell S, *et al.* (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, **26**,  
851 229-246.

852 Adjeroud M, Guérécheau A, Vidal-Dupiol J, Flot J. F, Arnaud-Haond S, Bonhomme F (2014) Genetic diversity,  
853 clonality and connectivity in the scleractinian coral *Pocillopora damicornis*: a multi-scale analysis in an  
854 insular, fragmented reef system. *Marine Biology*, **161**, 531-541.

855 Alberto F, Massa S, Manent P *et al.* (2008) Genetic differentiation and secondary contact zone in the  
856 seagrass *Cymodocea nodosa* across the Mediterranean-Atlantic transition region. *Journal of*  
857 *Biogeography*, **35**, 1279-1294.

858 Appeltans W, Ahyong ST, Anderson G *et al.* (2012) The magnitude of global marine species diversity. *Current*  
859 *Biology*, **22**, 2189-2202.

860 Arnaud-Haond S, Alberto F, Teixeira S, Procaccini G, Serrao EA, Duarte CM (2005) Assessing genetic diversity  
861 in clonal organisms: Low diversity or low resolution? Combining power and cost efficiency in selecting  
862 markers. *Journal of Heredity*, **96**, 434-440.

863 Arnaud-Haond S, Duarte CM, Alberto F, Serrao EA (2007a) Standardizing methods to address clonality in  
864 population studies. *Molecular Ecology*, **16**, 5115-5139.

865 Arnaud-Haond S, Duarte CM, Diaz-Almela E, Marbà N, Sintes T, Serrão, EA (2012) Implications of Extreme  
866 Life Span in Clonal Organisms: Millenary Clones in Meadows of the Threatened Seagrass *Posidonia*  
867 *oceanica*. *PLoS ONE*, **7**, e30454.

868 Arnaud-Haond S, Migliaccio M, Diaz-Almela E *et al.* (2007b) Vicariance patterns in the Mediterranean Sea:  
869 east-west cleavage and low dispersal in the endemic seagrass *Posidonia oceanica*. *Journal of*  
870 *Biogeography*, **34**, 963-976.

871 Arnaud-Haond S, Moalic Y, Hernández-García E *et al.* (2014) Disentangling the Influence of Mutation and  
872 Migration in Clonal Seagrasses Using the Genetic Diversity Spectrum for Microsatellites. *Journal of*  
873 *Heredity*, **105**, 532-541.

874 Avise J (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, **36**, 3-15.

875 Baird NA, Etter PD, Atwood TS *et al.* (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced  
876 RAD Markers. *PLoS ONE*, **3**, e3376.

877 Baker CS, Steel D, Calambokidis J *et al.* (2013) Strong maternal fidelity and natal philopatry shape genetic  
878 structure in North Pacific humpback whales. *Marine Ecology Progress Series*, **494**, 291-306.

879 Barberousse A, Samadi S (2010) Species from Darwin onward. *Integrative Zoology*, **5**, 187-197.

- 880 Becheler R, Benkara E, Moalic Y, Hily C, Arnaud-Haond S (2014) Scaling of processes shaping the clonal  
881 dynamics and genetic mosaic of seagrasses through temporal genetic monitoring. *Heredity*, **112**, 114-  
882 121.
- 883 Becheler R, Diekmann O, Hily C, Moalic Y, Arnaud-Haond S (2010) The concept of population in clonal  
884 organisms: mosaics of temporally colonized patches are forming highly diverse meadows of *Zostera*  
885 *marina* in Brittany. *Molecular Ecology*, **19**, 2394-2407.
- 886 Billard E, Daguin C, Pearson G, Serrão E, Engel C, Valero M (2005) Genetic isolation between the three  
887 closely related taxa *Fucus vesiculosus*, *F. spiralis* and *F. ceranoides*. *Journal of Phycology*, **41**, 900-905.
- 888 Billard E, Serrão E, Pearson G, Destombe C, Valero M (2010) *Fucus vesiculosus* and *spiralis* species complex:  
889 a nested model of local adaptation at the shore level. *Marine Ecology Progress Series*, **405**, 163-174.
- 890 Bilton DT, Paula J, Bishop JDD (2002) Dispersal, genetic differentiation and speciation in estuarine  
891 organisms. *Estuarine, Coastal and Shelf Science*, **55**, 937-952.
- 892 Bishop JDD, Roby C, Yunnice ALE *et al.* (2013) The Southern Hemisphere ascidian *Asterocarpa humilis* is  
893 unrecognised but widely established in NW France and Great Britain. *Biological Invasions*, **15**, 253-  
894 260.
- 895 Blackburn TM, Pysek P, Bacher S, Carlton JT, Duncan RP, Jarosík V, Wilson JRU, Richardson DM (2011) A  
896 proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, **26**, 333-339.
- 897 Bock DG, MacIsaac HJ, Cristescu ME (2012) Multilocus genetic analyses differentiate between widespread  
898 and spatially restricted cryptic species in a model ascidian. *Proceedings of the Royal Society B:*  
899 *Biological Sciences*, **279**, 2377-2385.
- 900 Bors EK, Rowden AA, Maas EW, Clark MR, Shank TM (2012) Patterns of deep-sea genetic connectivity in the  
901 New Zealand region: implications for management of benthic ecosystems. *PLoS ONE*, **7**, e49474.
- 902 Bortolus A (2008) Error cascades in the biological sciences: the unwanted consequences of using bad  
903 taxonomy in ecology. *AMBIO: A Journal of the Human Environment*, **37**, 114-118.
- 904 Bottom DL, Jones KK, Cornwell TJ, Gray A, Simenstad CA (2005) Patterns of Chinook salmon migration and  
905 residency in the Salmon River estuary (Oregon). *Estuarine Coastal and Shelf Science*, **64**, 79-93.
- 906 Bucklin A, Steinke D, Blanco-Bercial L (2011) DNA barcoding of marine metazoa. *Annual Review of Marine*  
907 *Science*, **3**, 471-508.
- 908 Calderón I, Garrabou J, Aurelle D (2006) Evaluation of the utility of COI and ITS markers as tools for  
909 population genetic studies of temperate gorgonians. *Journal of Experimental Marine Biology and*  
910 *Ecology*, **336**, 184-197.
- 911 Camargo A, Sites J (2013) Species delimitation: a decade after the Renaissance. In: The species problem:  
912 ongoing issues (ed Pavlinov IY), pp 225-247. InTech – Open Access publisher.



- 913 Caputi L, Andreakis N, Mastrototaro F, Cirino P, Vassillo M, Sordino P (2007) Cryptic speciation in a model  
914 invertebrate chordate. *Proceedings of the National Academy of Sciences*, **104**, 9364-9369.
- 915 Carlton J T (2009) Deep invasion ecology and the assembly of communities in historical time. In: Biological  
916 invasions in marine ecosystems (eds Rilov G, Crooks JA), pp 13-56. Springer, Berlin Heidelberg.
- 917 Carlton JT (1996) Biological invasions and cryptogenic species. *Ecology*, **77**, 1653-1655.
- 918 Carstens BC, Pelletier TA, Reid NM, Satler JD (2013) How to fail at species delimitation. *Molecular Ecology*,  
919 **22**, 4369-4383.
- 920 Castelin M, Lambourdière J, Boisselier MC *et al.* (2010) Hidden diversity and endemism on seamounts: focus  
921 on poorly dispersive neogastropods. *Biological Journal of the Linnean Society*, **100**, 420-438.
- 922 Castelin M, Lorion J, Brisset J *et al.* (2012) Speciation patterns in gastropods with long-lived larvae from  
923 deep-sea seamounts. *Molecular ecology*, **21**, 4828-4853.
- 924 Chenuil A, Hoareau TB, Egea E *et al.* (2010) An efficient method to find potentially universal population  
925 genetic markers, applied to metazoans. *BMC Evolutionary Biology*, **10**, 276.
- 926 Collins RA, Cruickshank RH (2013) The seven deadly sins of DNA barcoding. *Molecular Ecology Resources*, **13**,  
927 969-975.
- 928 Costello MJ, Coll M, Danovaro R, Haplin P, Ojaveer H, Miloslavich P (2010) A census of marine biodiversity  
929 knowledge, resources, and future challenges. *PLoS ONE*, **5**, e11842.
- 930 Cowen RK, Lwiza KM, Sponaugle S, Paris CB, Olson DB (2000) Connectivity of marine populations: open or  
931 closed? *Science*, **287**, 857-859.
- 932 Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. *Annual Review of*  
933 *Marine Science*, **1**, 443-466.
- 934 Coyer JA, Diekmann OE, Serrao EA *et al.* (2004) Population genetics of dwarf eelgrass *Zostera nolti*  
935 throughout its biogeographic range. *Marine Ecology-Progress Series*, **281**, 51-62.
- 936 Coyer JA, Hoarau G, Costa J F *et al.* (2011). Evolution and diversification within the intertidal brown  
937 macroalgae *Fucus spiralis* / *F. vesiculosus* species complex in the North Atlantic. *Molecular*  
938 *Phylogenetics and Evolution*, **58**, 283-296.
- 939 Coyne JA, Orr HA (2004) *Speciation*. Sunderland, MA: Sinauer Associates. 545 pp.
- 940 Dayrat B (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society*, **85**, 407-415.
- 941 de Queiroz K (1998) The general concept of species, species criteria, and the process of speciation. In:  
942 Endless forms: species and speciation (eds Howard D J, Berlocher SH), pp 57-75. Oxford University  
943 Press, Oxford.
- 944 Dehal P, Satou Y, Campbell RK *et al.* (2002) The draft genome of *Ciona intestinalis*: Insights into chordate  
945 and vertebrate origins. *Science*, **298**, 2157-2167.

- 946 DeSalle R, Egan MG, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding  
947 . *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**, 1905-1916.
- 948 Destombe C, Valero M, Guillemin M-L (2010) Delineation of two sibling red algal species, *Gracilaria gracilis*  
949 and *Gracilaria dura* (Gracilariales, Rhodophyta), using multiple DNA markers: resurrection of the  
950 species *G. dura* previously described in the northern Atlantic 200 years ago. *Journal of Phycology*, **46**,  
951 720-727.
- 952 Dijoux L, Viard F, Payri C (2014) The more we search, the more we find: discovery of a new lineage and a  
953 new species complex in the genus *Asparagopsi*. *PLoS ONE*, **9**, e103826.
- 954 Elton CS (1958). *The ecology of invasion by animals and plants*. 181p. London, Methuen.
- 955 Engel C, Destombe C, Valero M (2004) Mating system and gene flow in the red seaweed *Gracilaria gracilis*:  
956 effect of haploid-diploid life history and intertidal rocky shore landscape on finescale genetic  
957 structure. *Heredity*, **92**, 289-298.
- 958 Engel CR, Daguin C, Serrao EA (2005) Genetic entities and mating system in hermaphroditic *Fucus spiralis*  
959 and its close dioecious relative *F. vesiculosus* (Fucaceae, Phaeophyceae). *Molecular ecology*, **14**, 2033-  
960 2046.
- 961 Engelhaupt D, Hoelzel AR, Nicholson C *et al.* (2009) Female philopatry in coastal basins and male dispersion  
962 across the North Atlantic in a highly mobile marine species, the sperm whale (*Physeter*  
963 *macrocephalus*). *Molecular Ecology*, **18**, 4193-4205.
- 964 Fogarty MJ, Botsford LW (2007) Population connectivity and spatial management of marine fisheries.  
965 *Oceanography*, **20**, 112-123.
- 966 Fontaine B, Perrard A, Bouchet P (2012) 21 years of shelf life between discovery and description of new  
967 species. *Current Biology*, **22**, R943-944.
- 968 Funk D, Omland K (2003) Species-level paraphyly and polyphyly: Frequency, causes and consequences, with  
969 insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution and Systematics*, **34**,  
970 397-423.
- 971 Gérard K, Guilloton E, Arnaud-Haond S *et al.* (2013) PCR survey of 50 introns in animals: Cross-amplification  
972 of homologous EPIC loci in eight non-bilaterian, protostome and deuterostome phyla. *Marine*  
973 *Genomics*, **12**, 1-8.
- 974 González A, Beltrán J, Hiriart-Bertrand L *et al.* (2012) Identification of cryptic species in the *Lessonia*  
975 *nigrescens* complex (Phaeophyceae, Laminariales). *Journal of Phycology*, **48**, 1153-1165.
- 976 Guillemin M-L, Faugeron S, Destombe C, Viard F, Correa JA, Valero M (2008) Genetic variation in wild and  
977 cultivated populations of the haploid-diploid red alga *Gracilaria chilensis*: how farming practices favor  
978 asexual reproduction and heterozygosity. *Evolution*, **62**, 1500-1519.

- 979 Halbert KM, Goetze E, Carlon DB (2013) High Cryptic Diversity across the Global Range of the Migratory  
980 Planktonic Copepods *Pleuromamma piseki* and *P. gracilis*. *PLoS ONE*, **8**, e77011.
- 981 Havermans C, Sonet G, d'Acoz CDU *et al.* (2013) Genetic and morphological divergences in the cosmopolitan  
982 deep-sea amphipod *Eurythenes gryllus* reveal a diverse abyss and a bipolar species. *PLoS ONE*, **8**,  
983 e74218.
- 984 Haydar D (2012) What is natural? The scale of cryptogenesis in the North Atlantic Ocean. *Diversity and*  
985 *Distributions*, **18**, 101-110.
- 986 Haydar D, Hoarau G, Olsen JL, Stam WT, Wolff WJ (2011) Introduced or glacial relict? Phylogeography of the  
987 cryptogenic tunicate *Molgula manhattensis* (Asciacea, Pleurogona). *Diversity and Distributions*, **17**,  
988 68-80.
- 989 Hsieh C-H, Ko C-C, Chung C-H, Wang H-Y (2014) Multilocus approach to clarify species status and the  
990 divergence history of the Bemisia tabaci (Hemiptera: Aleyrodidae) species complex. *Molecular*  
991 *Phylogenetics and Evolution*, **76**, 172-80
- 992 Jaafar TNAM, Taylor MI, Nor SAM, de Bruyn M, Carvalho GR (2012) DNA Barcoding Reveals Cryptic Diversity  
993 within Commercially Exploited Indo-Malay Carangidae (Teleostei: Perciformes). *PLoS ONE*, **7**, e49623.
- 994 Jarman S, Ward R, Elliott N (2002) Oligonucleotide primers for PCR amplification of coelomate introns.  
995 *Marine Biotechnology*, **4**, 347-355.
- 996 Johnson GD, Paxton JR, Sutton TT *et al.* (2009) Deep-sea mystery solved: astonishing larval transformations  
997 and extreme sexual dimorphisms unite three fish families. *Biology Letters*, **5**, 235-239.
- 998 Jones GP, Srinivasan M, Almany GR (2007) Population connectivity and conservation of marine biodiversity.  
999 *Oceanography*, **20**, 100-111.
- 1000 Kekkonen M, Hebert PD (2014) DNA barcode-based delineation of putative species: efficient start for  
1001 taxonomic workflows. *Molecular ecology resources*, **14**, 706-715.
- 1002 Kendrick GA, Waycott M, Carruthers TJB *et al.* (2012) The Central Role of Dispersal in the Maintenance and  
1003 Persistence of Seagrass Populations. *Bioscience*, **62**, 56-65.
- 1004 Kinlan BP, Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: a community  
1005 perspective. *Ecology*, **84**, 2007-20.
- 1006 Knowlton N (2000) Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia*, **420**, 73-90.
- 1007 Komai T, Segonzac M (2005) A revision of the genus *Alvinocaris* Williams and Chace (Crustacea: Decapoda:  
1008 Caridea: Alvinocarididae), with descriptions of a new genus and a new species of *Alvinocaris*. *Journal*  
1009 *of Natural History*, **39**, 1111-1175.
- 1010 Kritzer JP, Sale PF (2004) Metapopulation ecology in the sea: from Levin's model to marine ecology and  
1011 fisheries science. *Fish and Fisheries*, **5**, 131-140.

- 1012 Krueger-Hadfield SA, Collén J, Daguin-Thiébaud C, Valero M (2011) Genetic population structure and mating  
1013 system in *Chondrus crispus* (Rhodophyta). *Journal of Phycology*, **47**, 440-450.
- 1014 Leliaert F, Verbruggen H, Vanormelingen P *et al.* (2014) DNA-based species delimitation in algae. *European*  
1015 *Journal of Phycology*, **49**, 179-196.
- 1016 Leaché AD, Fujita MK, Minin VN, Bouckaert RR (2014). Species delimitation using genome-wide SNP data.  
1017 *Systematic Biology*, **63**, 534-542.
- 1018 Lexer C, Wüest RO, Mangii S *et al.* (2014) Genomics of the divergence continuum in an African plant  
1019 biodiversity hotspot, I: drivers of population divergence in *Restio capensis* (Restionaceae). *Molecular*  
1020 *Ecology*, **23**, 4373-4386.
- 1021 López-Cristoffanini C, Tellier F, Otaíza R, Correa Juan A, Contreras-Porcía L (2013) Tolerance to air exposure:  
1022 a feature driving the latitudinal distribution of two sibling kelp species. *Botanica Marina*, **56**, 431-440.
- 1023 Mantelatto FL, Pezzuto PR, Masello A, Rossi Wongtschowski CLDB, Silva Hilsdorf AW, Rossi N (2014)  
1024 Molecular analysis of the commercial deep-sea crabs *Chaceon ramosae* and *Chaceon notialis*  
1025 (Brachyura, Geryonidae) reveals possible cryptic species in the South Atlantic. *Deep Sea Research Part*  
1026 *I: Oceanographic Research Papers*, **84**, 29-37.
- 1027 McFadden CS, Benayahu Y, Pante E, Thoma JN, Nevarez PA, France SC (2011) Limitations of mitochondrial  
1028 gene barcoding in Octocorallia. *Molecular Ecology Resources*, **11**, 19-31.
- 1029 McFadden CS, Sánchez JA, France SC (2010) Molecular phylogenetic insights into the evolution of  
1030 Octocorallia: A review. *Integrative and Comparative Biology*, **50**, 389-410.
- 1031 McGlashan DJ, Ponniah M, Cassey P, Viard F (2008) Clarifying marine invasions with molecular markers: an  
1032 illustration based on mtDNA from mistaken calyptraeid gastropod identifications. *Biological Invasions*,  
1033 **10**, 51-57.
- 1034 Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. *PLoS biology*, **3**,  
1035 e422.
- 1036 Miller K, Williams A, Rowden AA, Knowles C, Dunshea G (2010) Conflicting estimates of connectivity among  
1037 deep-sea coral populations. *Marine Ecology*, **31**, 144-157.
- 1038 Mishler BD (2009) Species are not uniquely real biological entities. In: Contemporary debates in philosophy  
1039 of biology (eds Ayala FJ, Arp R), pp 110-122. Wiley-Blackwell, Oxford UK.
- 1040 Moritz C (1994) Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular*  
1041 *Ecology*, **3**, 401-411.
- 1042 Nielsen EE, Hemmer-Hansen J, Poulsen NA *et al.* (2009) Genomic signatures of local directional selection in a  
1043 high gene flow marine organism ; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology*, **9**, 276.
- 1044 Not F, Siano R, Kooistra WH *et al.* (2012) Diversity and Ecology of Eukaryotic Marine Phytoplankton.  
1045 *Advances in Botanical Research*, **64**, 1-53.

- 1046 Nydam M, Harrison R (2007) Genealogical relationships within and among shallow-water *Ciona* species  
1047 (Asciacea). *Marine Biology*, **151**, 1839-1847.
- 1048 Nydam ML, Harrison RG (2011) Introgression despite substantial divergence in a broadcast spawning marine  
1049 invertebrate. *Evolution*, **65**, 429-442.
- 1050 Olsen JL, Stam WT, Coyer JA *et al.* (2004) North Atlantic phylogeography and large-scale population  
1051 differentiation of the seagrass *Zostera marina* L. *Molecular Ecology*, **13**, 1923-1941.
- 1052 Oppliger LV, Correa JA, Engelen AH *et al.* (2012) Temperature Effects on Gametophyte Life-History Traits  
1053 and Geographic Distribution of Two Cryptic Kelp Species. *PLoS ONE*, **7**, e39289.
- 1054 Padial JM, Miralles A, De la Riva I, Vences M (2010) Review: The integrative future of taxonomy. *Frontiers in*  
1055 *Zoology*, **7**, 1-14.
- 1056 Palsbøll PJ, Bérubé M, Allendorf F (2007) Identification of management units using population genetic data.  
1057 *Trends in Ecology and Evolution*, **22**, 11-16.
- 1058 Pante E, Abdelkrim J, Viricel A, Gey D, France S, Boisselier MC, Samadi S (accepted) Use of RAD sequencing  
1059 for delimiting species. *Heredity*
- 1060 Plouviez S, Schultz T, McGinnis G, Minshall H, Rudder M, Van Dover CL (2013) Genetic diversity of  
1061 hydrothermal-vent barnacles in Manus Basin. *Deep Sea Research Part I: Oceanographic Research*  
1062 *Papers*, **82**, 73-79.
- 1063 Pons J, Barraclough TG, Gomez-Zurita J, *et al.* (2006) Sequence based species delimitation for the DNA  
1064 taxonomy of undescribed insects. *Systematic Biology*, **55**, 595-609.
- 1065 Puillandre N, Lambert A, Brouillet S, Achaz G (2012b) ABGD, Automatic Barcode Gap Discovery for primary  
1066 species delimitation. *Molecular Ecology*, **21**, 1864-1877.
- 1067 Puillandre N, Macpherson E, Lambourdière J, Cruaud C, Boisselier-Dubayle MC, Samadi S (2011) Barcoding  
1068 type specimens helps to identify synonyms and an unnamed new species in *Eumunida* Smith, 1883  
1069 (Decapoda: Eumunididae). *Invertebrate Systematics*, **25**, 322-333.
- 1070 Puillandre N, Modica MV, Zhang Y *et al.* (2012a) Large-scale species delimitation method for hyperdiverse  
1071 groups. *Molecular ecology*, **21**, 2671-2691.
- 1072 Reiss H, Hoarau G, Dickey-Collas M, Wolff WJ (2009) Genetic population structure of marine fish: mismatch  
1073 between biological and fisheries management units. *Fish and Fisheries*, **10**, 361-395.
- 1074 Riginos C, Douglas KE, Jin Y, Shanahan DF, Trembl EA (2011) Effects of geography and life history traits on  
1075 genetic differentiation in benthic marine fishes. *Ecography*, **34**, 566-75.
- 1076 Rilov G, Crooks Eds. (2009) *Biological Invasions in Marine Ecosystems: Ecological, Management, and*  
1077 *Geographic Perspectives*. Ecological Studies. 641 pages. Springer.

- 1078 Ritt B, Duperron S, Lorion J, Sara Lazar C, Sarrazin J (2012) Integrative study of a new cold-seep mussel  
1079 (Mollusca: Bivalvia) associated with chemosynthetic symbionts in the Marmara Sea. *Deep Sea*  
1080 *Research Part I: Oceanographic Research Papers*, **67**, 121-132.
- 1081 Rocha LA, Robertson DR, Roman J, Bowen BW (2005) Ecological speciation in tropical reef fishes.  
1082 *Proceedings of the Royal Society B*, **272**, 573-579.
- 1083 Rocha LA, Rocha CR, Robertson DR, Bowen BW (2008) Comparative phylogeography of Atlantic reef fishes  
1084 indicates both origin and accumulation of diversity in the Caribbean. *BMC Evolutionary Biology*, **8**,  
1085 157.
- 1086 Roux C, Tasgkogeorga G, Bierne N, Galtier N (2013) Crossing the Species Barrier: Genomic Hotspots of  
1087 Introgression between Two Highly Divergent *Ciona intestinalis* Species. *Molecular Biology and*  
1088 *Evolution*, **30**, 1574-1587.
- 1089 Rozenfeld AF, Arnaud-Haond S, Hernández-García E *et al.* (2007) Spectrum of genetic diversity and networks  
1090 of clonal populations. *Journal of the Royal Society Interface*, **4**, 1093-1102.
- 1091 Samadi S, Barberousse A (2006) The tree, the network, and the species. *Biological Journal of the Linnean*  
1092 *Society*, **89**, 509-521.
- 1093 Samadi S, Barberousse A. (2014). Species. *In press*. In: Handbook of Evolution Thinking in the Sciences  
1094 (Heams, Huneman, Lecointre, Silberstein, eds). Springer.
- 1095 Santelices B (1989) *Algas marinas de Chile*. 399 pp. Ediciones Universidad Católica de Chile, Santiago, Chile.
- 1096 Satler JD, Carstens BC, Hedin M (2013) Multilocus Species Delimitation in a Complex of Morphologically  
1097 Conserved Trapdoor Spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*). *Systematic Biology* **62**,  
1098 805-823.
- 1099 Selkoe KA, Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics  
1100 of dispersal. *Marine Ecology Progress Series*, **436**, 291-305 .
- 1101 Serrão EA, Alice LA, Brawley SH (1999) Evolution of the Fucaceae (Phaeophyceae) inferred from nrDNA-ITS.  
1102 *Journal of Phycology*, **35**, 382-394.
- 1103 Shank TM (2010) Seamounts: Deep-oceans laboratories of faunal connectivity, evolution, and endemism.  
1104 *Oceanography*, **23**, 108-122.
- 1105 Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. *Biological Bulletin*, **216**, 373-85.
- 1106 Shearer TL, Coffroth MA (2008) Barcoding corals: limited by interspecific divergence, not intraspecific  
1107 variation. *Molecular Ecology Resources*, **8**, 247-255.
- 1108 Siegel DA, Kinlan BP, Gaylord B, Gaines SD (2003) Lagrangian descriptions of marine larval dispersion.  
1109 *Marine Ecology Progress Series*, **260**, 83-96.

- 1110 Simberloff D, Martin J-L, Genovesi P, Maris V, Wardle DA, Aronson J, Courchamp F, Galil B, García-Berthou E,  
1111 Pascal M, Pyšek P, Sousa R, Tabacchi E, Vilà M (2013) Impacts of biological invasions: what's what  
1112 and the way forward. *Trends in Ecology & Evolution*, **28**, 58-66.
- 1113 Sosa P, Valero M, Batista F, Gonzalez-Perez M (1998) Genetic structure of natural populations of *Gelidium*  
1114 species: A re-evaluation of results. *Journal of Applied Phycology*, **10**, 279-284.
- 1115 Sosa PA, Garcia-Reina G (1993) Genetic variability of *Gelidium canariensis* (Rhodophyta) determined by  
1116 isozyme electrophoresis. *Journal of Phycology*, **29**, 118-124.
- 1117 Suzuki MM, Nishikawa T, Bird A (2005) Genomic approaches reveal unexpected genetic divergence within  
1118 *Ciona intestinalis*. *Journal of Molecular Evolution*, **61**, 627-635.
- 1119 Taylor MS, Hellberg ME (2003) Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish.  
1120 *Science*, **299**, 107-109.
- 1121 Taylor MS, Hellberg ME (2005) Marine radiations at small geographical scales: speciation in Neotropical reef  
1122 gobies (*Elacatinus*). *Evolution*, **59**, 374-385.
- 1123 Teixeira S, Olu K, Decker C *et al.* (2013) High connectivity across the fragmented chemosynthetic ecosystems  
1124 of the deep Atlantic Equatorial Belt: efficient dispersal mechanisms or questionable endemism?  
1125 *Molecular Ecology*, **22**, 4663-4680.
- 1126 Tellier F, Meynard A, Correa J, Faugeron S, Valero M (2009) Phylogeographic analyses of the 30 degrees S  
1127 south-east Pacific biogeographic transition zone establish the occurrence of a sharp genetic  
1128 discontinuity in the kelp *Lessonia nigrescens*: Vicariance or parapatry? *Molecular Phylogenetics and*  
1129 *Evolution*, **53**, 679-693.
- 1130 Tellier F, Tapia J, Faugeron S, Destombe C, Valero M (2011) The *Lessonia nigrescens* species complex  
1131 (Laminariales, Phaeophyceae) shows strict parapatry and complete reproductive isolation in a  
1132 secondary contact zone. *Journal of Phycology*, **47**, 894-903.
- 1133 Thomas Jr RC, Willette DA, Carpenter KE, Santos MD (2014) Hidden Diversity in Sardines: Genetic and  
1134 Morphological Evidence for Cryptic Species in the Goldstripe Sardinella, *Sardinella gibbosa* (Bleeker,  
1135 1849). *PLoS ONE*, **9**, e84719.
- 1136 Thubaut J, Puillandre N, Faure B, Cruaud C, Samadi S (2013) The contrasted evolutionary fates of deep-sea  
1137 chemosynthetic mussels (Bivalvia, Bathymodiolinae). *Ecology and evolution*, **3**, 4748-4766.
- 1138 Uthicke S, Byrne M, Conand C (2010) Genetic barcoding of commercial Bêche de mer species  
1139 (Echinodermata: Holothuroidea). *Molecular Ecology Resources*, **10**, 634-646.
- 1140 van der Strate HJ, van de Zande L, Stam WT, Olsen JL (2002) The contribution of haploids, diploids and  
1141 clones to fine-scale population structure in the seaweed *Cladophoropsis membranacea* (Chlorophyta).  
1142 *Molecular Ecology*, **11**, 329-345.

- 1143 Wahl M, Jormalainen V, Eriksson BK *et al.* (2011) Stress ecology in *Fucus*: abiotic, biotic and genetic  
1144 interactions. *Advances in marine biology*, **59**, 37-106.
- 1145 Wang JY, Frasier TR, Yang SC, White BN (2008) Detecting recent speciation events: the case of the finless  
1146 porpoise (genus *Neophocaena*). *Heredity*, **101**, 145-155.
- 1147 Waples RS, Punt AE, Cope JM (2008) Integrating genetic data into management of marine resources: how  
1148 can we do it better?. *Fish and Fisheries*, **9**, 423-449.
- 1149 Webb TJ, Berghe EV, O'Dor R (2010) Biodiversity's big wet secret: the global distribution of marine biological  
1150 records reveals chronic under-exploration of the deep pelagic ocean. *PLoS ONE*, **5**, e10223.
- 1151 Weersing K, Toonen RJ (2009) Population genetics, larval dispersal, and connectivity in marine systems.  
1152 *Marine Ecology Progress Series*, **393**, 12.
- 1153 Will KW, Mishler BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative  
1154 taxonomy. *Systematic biology*, **54**, 844-851.
- 1155 Williams AB (1988) New marine decapod crustaceans from waters influenced by hydrothermal, discharge,  
1156 brine and hydrocarbon seepage. *Fishery Bulletin*, **86**, 263–287.
- 1157 Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci U*  
1158 *S A*, **107**, 9264–9.
- 1159 Yeates DK, Seago A, Nelson L, Cameron SL, Joseph L, Trueman JW (2011) Integrative taxonomy, or iterative  
1160 taxonomy?. *Systematic Entomology*, **36**, 209-217.
- 1161 Zardi GI, Nicastro KR, Canovas F, Costa JF, Serrao EA, Pearson GA (2011) Adaptive Traits Are Maintained on  
1162 Steep Selective Gradients despite Gene Flow and Hybridization in the Intertidal Zone. *PLoS ONE*, **6**,  
1163 e19402.
- 1164 Zhan A, MacIsaac HJ, Cristescu ME (2010) Invasion genetics of the *Ciona intestinalis* species complex: from  
1165 regional endemism to global homogeneity. *Molecular Ecology*, **19**, 4678-4694.
- 1166