



**HAL**  
open science

## Molecular species delimitation methods and population genetics data reveal extensive lineage diversity and cryptic species in Aglaopheniidae (Hydrozoa)

Bautisse Postaire, H el ene Magalon, Chlo e A.-F. Annie-France Bourmaud, J. Henrich Bruggemann

### ► To cite this version:

Bautisse Postaire, H el ene Magalon, Chlo e A.-F. Annie-France Bourmaud, J. Henrich Bruggemann. Molecular species delimitation methods and population genetics data reveal extensive lineage diversity and cryptic species in Aglaopheniidae (Hydrozoa). *Molecular Phylogenetics and Evolution*, 2016, 105, pp.36-49. 10.1016/j.ympev.2016.08.013 . hal-01359407

**HAL Id: hal-01359407**

**<https://hal.science/hal-01359407>**

Submitted on 2 Sep 2016

**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 destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

# Molecular species delimitation methods and population genetics data reveal extensive lineage diversity and cryptic species in Aglaopheniidae (Hydrozoa)

Bautisse Postaire<sup>a,b,c,\*</sup>, H el ene Magalon<sup>a,b</sup>, Chlo e A.-F. Bourmaud<sup>a,b</sup>, J. Henrich Bruggemann<sup>a,b</sup>

<sup>a</sup>UMR ENTROPIE Universit e de La R eunion-CNRS-IRD, UFR Sciences et Technologies, 15 Avenue Ren e Cassin, CS 92003, 97744 Saint Denis, France

<sup>b</sup>Laboratoire d'Excellence CORAIL, France<sup>1</sup>

<sup>c</sup>IMBE UMR 7263, Aix Marseille Universit e-CNRS-IRD-Avignon Universit e, Station marine d'Endoume, Chemin de la batterie des lions, 13007 Marseille, France.

## A B S T R A C T

A comprehensive inventory of global biodiversity would be greatly improved by automating methods for species delimitation. The Automatic Barcode Gap Discovery method, the Poisson tree processes algorithm and the Generalized mixed Yule coalescent model have been proposed as means of increasing the rate of biodiversity description using single locus data. We applied these methods to explore the diversity within the Aglaopheniidae, a hydrozoan family with many species widely distributed across tropical and temperate oceans. Our analyses revealed widespread cryptic diversity in this family, almost half of the morpho species presenting several independent evolutionary lineages, as well as support for cases of synonymy. For two common species of this family, *Lytocarpia brevirostris* and *Macrorhynchia phoenicea*, we compared the outputs to clustering analyses based on microsatellite data and to nuclear gene phylogenies. For *L. brevirostris*, microsatellite data were congruent with results of the species delimitation methods, revealing the existence of two cryptic species with Indo Pacific distribution. For *M. phoenicea*, all analyses confirmed the presence of two cryptic species within the South Western Indian Ocean. Our study suggests that the diversity of Aglaopheniidae might be much higher than assumed, likely related to low dispersal capacities. Sequence based species delimitation methods seem highly valuable to reveal cryptic diversity in hydrozoans; their application in an integrative framework will be very useful in describing the phyletic diversity of these organisms.

## 1. Introduction

The global biodiversity crisis requires focusing conservation efforts on key areas that ensure the long term persistence of the greatest fraction of global biodiversity (Barnosky et al., 2011; Myers et al., 2000). The identification of such regions requires reliable assessments of alpha diversity, i.e. the number of species present in the area of interest (Margules and Pressey, 2000). Historically, species have been described and identified based on morphological characters. Even if traditional taxonomic work continues to be important for species inventories and conservation (Daugherty et al., 1990), ignoring cryptic diversity leads to incomplete taxon sampling and erroneous assessments of biodiversity, biogeographic patterns and speciation processes (Heath et al., 2008). Indeed, traditional taxonomy does not integrate genetic

diversity and might ignore potential cryptic species [i.e. two distinct sympatric species classified under the same taxonomic name based on morphological characters (see box 1 in Bickford et al., 2007)]. Hence, protecting only the 'visible biodiversity' will negatively affect conservation and management efforts of biodiversity and evolutionary processes (Briggs, 2005; Moritz, 2002, 1999).

Most evolutionary biologists concur that species and higher taxonomic levels form independent genealogical lineages of organisms (Mayden, 2002; Samadi and Barberousse, 2006) and the broad use of molecular markers during the past decades has revealed the prevalence of cryptic lineage diversity among marine organisms, especially in marine invertebrates (Boissin et al., 2008; Duda et al., 2008; Hoareau et al., 2013; Huelsken et al., 2013; Knowlton, 1993; Lindner et al., 2011; Niemiller et al., 2011; Palumbi, 1994; Pfenninger and Schwenk, 2007; Prada et al., 2014). The main issue in species delimitation is not the definition of what a species is, but resides in the issue of choosing a criterion (e.g. morphology, ecology, genetic distances, etc.) to identify lineages. No consensus exists (De Queiroz, 2007, 2005, 1992) and

\* Corresponding author.

E-mail address: bautisse.postaire@imbe.fr (B. Postaire).

<sup>1</sup> www.labex-coral.fr.

the criterion used can possibly lead to a false representation of a taxon's evolution and diversity (Agapow et al., 2004; De Queiroz, 2005; Mayden, 2002; Samadi and Barberousse, 2006). Thus, when describing alpha diversity, identification and description of species using morphological characters generate primary species hypotheses (PSHs) that need to be tested and confronted to other sources of information in order to delineate robust secondary species hypotheses (SSHs) and potentially reveal cryptic species [see Puillandre et al. (2012b) for an example in gastropods].

Among marine invertebrates, cnidarians in particular can be expected to present cryptic diversity due to the paucity of morphological characters useful for species description and systematics (e.g. Addamo et al., 2012; McFadden et al., 2014; Stampar et al., 2012). These morphological clues do not necessarily represent phylogenetic relationships since evolution of reproductive, ecological and physiological traits and eventually speciation do not always have morphological outcomes. Among cnidarians, the class Hydrozoa is particularly subject to morphological plasticity and taxonomic incertitude (Bavestrello et al., 2000; Bouillon et al., 2006; Leclère et al., 2009; Meroz Fine et al., 2003; Miglietta et al., 2009). Hydrozoans are found in almost all aquatic ecosystems: polar to tropical regions, shallow waters to abyssal plains, freshwater and marine ecosystems (Bouillon et al., 2006). Recent phylogenies showed that morphological characters thought to be taxonomically significant in this class were actually highly labile and plastic (e.g. Leclère et al., 2007; Miglietta et al., 2009; Moura et al., 2012; Postaire et al., 2015c). Several studies investigated cryptic diversity in hydrozoans (e.g. Schuchert, 2014; Folino Rorem et al., 2009; Moura et al., 2008; Govindarajan et al., 2005), and the existence of "true" cryptic species (*sensu* Bickford et al., 2007) in cosmopolitan morpho species (i.e. species delimited using morphological characters) has already been revealed or hypothesized (Lindner et al., 2011; Miglietta et al., 2007; Schuchert, 2005). Unfortunately, recent studies on hydrozoan systematics are often limited to a DNA barcoding approach, a method which was originally developed to help species identification by associating DNA sequences to type specimens (Hebert et al., 2003; Puillandre et al., 2011; Vernoooy et al., 2010). While DNA barcoding and genetic data in general were not initially proposed to be employed as species delimitation tools, it can help to uncover species diversity in complex taxa (Hebert et al., 2004) and several novel methods have been developed to use genetic data as a first step to delimit putative species, especially in taxa with limited comprehensive information (Carstens et al., 2013). These methods are particularly useful in taxa lacking clear synapomorphies and species boundaries, such as hydrozoans [see Castelin and Lambourdière (2010) for an example in gastropods].

Here we used three of these methods on the Aglaopheniidae (Agassiz, 1862), a highly specious family of mostly brooding hydrozoans with unresolved taxonomy and phylogeny (Bouillon et al., 2006; Moura et al., 2012; Postaire et al., 2015c). First, the Automatic Barcode Gap Discovery (ABGD) method (Puillandre et al., 2012a) uses a genetic distance based approach to detect a barcode gap dividing candidate species in the dataset by assuring that intra and inter specific genetic distances do not overlap. This method is independent of tree topology. ABGD calculates all pairwise distances and orders them as ranked values. A sliding window is then applied to calculate a local slope function across these values: the barcode gap is the first significant peak (increase of slope) that allows inferring primary partitions of the dataset. Each subpartition is then analyzed using the same approach until no further significant gaps are found (Puillandre et al., 2012a). Secondly, species delimitation methods based on the Generalized mixed Yule coalescent (GMYC) model (Fontaneto et al., 2010; Fujita et al., 2012; Pons et al., 2006) are based on tree topologies to infer species hypotheses. Using a likelihood function modeling evolutionary

processes, this model states that each node of a phylogenetic tree corresponds to one of two possible events: divergence between species following a strict Yule process [no extinction; (Yule, 1925)] or neutral coalescent events between lineages forming a species (Kingman, 1982). As coalescent events are assumed to occur at higher rates than speciation, it is thus possible to identify a limit on a phylogenetic tree between inter and intra specific divergence, delimiting clusters of leaves. Such clusters represent genetically isolated, independently evolving lineages, in which selection and genetic drift operate (Fujita et al., 2012), i.e. species hypotheses. Finally, the Poisson tree processes (PTP) species delimitation method is based on the differences between sequences (number of substitutions), but contrary to the GMYC models, it does not use a calibrated tree (Zhang et al., 2013). This method makes the assumptions that each mutation event has a non null probability of forming a new species and, as a consequence, that the number of substitutions between species is significantly higher than the number of substitutions within species. We used a combination of the three methods (i.e. ABGD, GMYC and PTP) to delineate SSHs.

Although considered efficient in identifying species limits (Puillandre et al., 2012b), even when singletons (i.e. a single haplotype per species hypothesis, PSH and/or SSH) represent an important part of the dataset (Talavera et al., 2013), some studies highlighted the tendency of these methods to overestimate the number of species (Hamilton et al., 2014; Lohse, 2009; Puillandre et al., 2012a; Zhang et al., 2013). Several studies further underlined that basing species delimitation solely on genetic data, a fortiori on a single mitochondrial marker, must be made cautiously (Dellicour and Flot, 2015; Hamilton et al., 2014; Jörger et al., 2012; Lohse, 2009). Indeed, deeply divergent mitochondrial lineages do not always imply distinct species: divergence might result from ancestral polymorphism, genetic introgression or hybridization (Ladner and Palumbi, 2012). Furthermore, gene trees do not always reflect species evolution (Hoelzer, 1997) and phylogenies must be discussed in an integrative framework that includes all available information (Dayrat, 2005; Padial et al., 2010; Puillandre et al., 2009; Schlick Steiner et al., 2010).

We tested SSHs for two PSHs, *Lytocarpia brevirostris* (Busk, 1852) and *Macrorhynchia phoenicea* (Busk, 1852), using newly developed microsatellite markers to assess the congruence between the ABGD method, GMYC models, the PTP method and population genetics data. Microsatellite data have already been efficiently used for this purpose (Hausdorf et al., 2011; Hausdorf and Hennig, 2010; Turini et al., 2014). Furthermore, microsatellite markers present the advantage of being codominant, biparentally inherited and neutral, unlike mitochondrial markers. Finally, we compared species delimitation outputs to phylogenies produced with two nuclear markers: the first exon of the calmodulin and a sequence comprising ITS1 and ITS2.

This study is the first to apply species delimitation methods based on DNA sequences in Aglaopheniidae, a highly diversified family lacking clear synapomorphies and/or species limits (Moura et al., 2012; Postaire et al., 2015c). It aims to clarify taxonomic issues associated with cryptic diversity by using a combination of population genetics data, molecular phylogenies and DNA based species delimitation methods.

## 2. Material and methods

### 2.1. Sampling of PSHs

The samples used for phylogenetic analyses are the same as in Postaire et al. (2015c). Between 2007 and 2014, we explored three marine provinces as defined by Spalding et al. (2007), comprising

six localities in two ecoregions of the Western Indian Ocean (WIO) province (western/northern Madagascar and Mascarene Islands), one locality in the South East Polynesian province (SEP; Moorea, Society Islands) and one in the tropical South Western Pacific province (SWP; New Caledonia) (Fig. 1a). We explored shallow coral reefs habitats (0–40 m depth) for Aglaopheniidae morpho species using SCUBA or snorkeling and picked samples manually. We preferentially collected parts of large colonies (i.e. feathers between 3 and 40 cm), referred herein to as individuals, at several decimeters distance to limit resampling the same genet (hydrozoans are capable of clonal propagation notably through stolonial growth). We identified individuals based on morphological characters used in routine for hydrozoans identification using the taxonomic literature [see references listed in Gravier Bonnet and Bourmaud (2012, 2006a, 2006b)]. We used the “Hydrozoa Handbook” (Bouillon et al., 2006) for species and genera lists. We aimed to collect and sequence a minimum of three individuals per morpho species, referred hereafter as primary species hypotheses (PSHs).

For population genetics analyses, we sampled new individuals of *L. brevisrostris* (n = 679, Fig. 1b) and *M. phoenicea* (n = 1337, Fig. 1c) following the same protocol as above at several locations in the WIO and in New Caledonia. Both PSHs were equally prospected in all locations, indicating that sampling discrepancies

represent the absence of the corresponding morpho species at the location. During sampling, we separated individuals of *M. phoenicea* into two groups, morpho types A and B, based on the general aspect of the colonies (general colony shape, color; Appendix A): morpho type A is usually tall (up to 10 cm), stiff and colored black and white, whereas morpho type B is more gracile and usually brown orange. The distinction was also based on empirical ecological data: both morpho types were found at similar depth but seemed to inhabit different microhabitats. Morpho type A occurs in well illuminated micro habitats exposed to strong currents, often at the base of *Pocillopora* sp. colonies, whereas morpho type B is found under sheltered cliff edges.

Specimens were fixed and preserved in 90% ethanol for DNA extraction. Whenever possible, a part of each sequenced individual was also preserved in a 3% formalin solution in seawater. Samples are available on request and stored in the biological collections of the Université de La Réunion.

## 2.2. DNA extraction

All reproductive structures were removed from each individual prior to DNA extraction. One or two ramifications per individual were used for extraction with DNeasy Blood & Tissue kit (Qiagen),

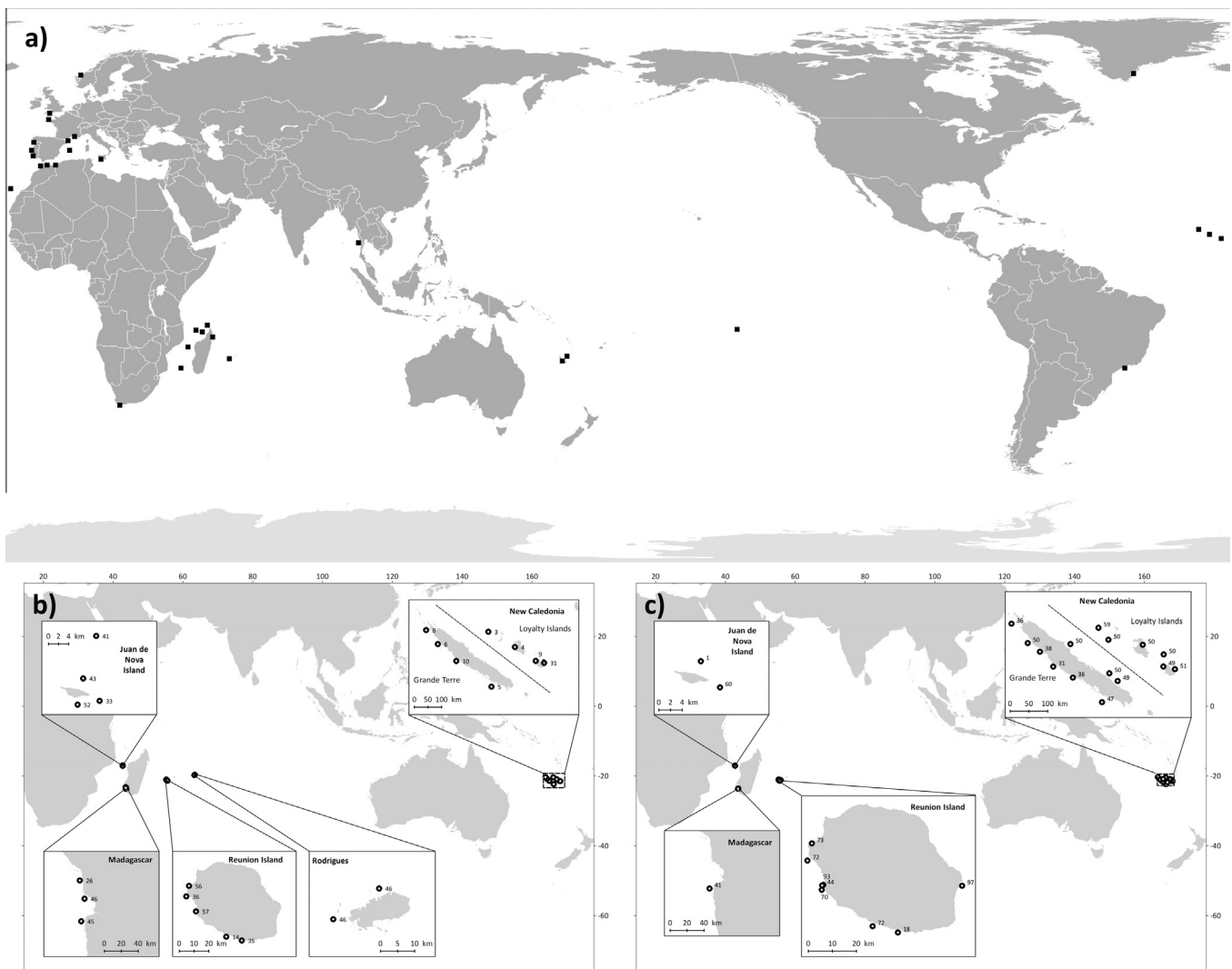


Fig. 1. (a) Main sampling sites of Aglaopheniidae specimens used in this study (Leclère et al., 2007, 2009; Moura et al., 2008, 2012; Postaire et al., 2015c) (b) sampling sites and number of samples of *Lytocarpia brevisrostris* (this study) and (c) *Macrorhynchia phoenicea* (this study).

following the manufacturer's protocol. We assessed the extraction quality on a 0.8% agarose gel stained with GelRed Nucleic Acid Stain, 10000X in DMSO (Gentaur).

### 2.3. Mitochondrial marker sequencing

We amplified and sequenced a fragment of the 16S rRNA (16S) mitochondrial gene with primers SHA (5' ACG GAA TGA ACT CAA ATC ATG T 3') and SHB (5' TCG ACT GTT TAC CAA AAA CAT A 3') (Cunningham and Buss, 1993). For *L. breviostris* and *M. phoenicea*, we also sequenced two nuclear regions for several individuals (representing main sampling sites and mitochondrial lineages): (1) a fragment of the first exon of the calmodulin (CAM) gene with primers CAMF1 (5' GAT CAA YTR CAN GAR GAA CAA ATT GC 3') and CAMR1 (5' CCA TCN CCA TCR ATA TCA GC 3') (Lindner et al., 2008), and (2) the complete internal transcribed spacer 1, 5.8S rDNA and internal transcribed spacer 2 region of the ribosomal DNA (ITS) with partial 18S and 28S ribosomal genes flanking both ITS, respectively upstream and downstream, with primers ITSF (5' CAC CGC CCG TCG CTA C TA CCG ATT GAA TGG 3') and ITS R (5' CGC TTC ACT CGC CGT TAC TAG GGG AAT CC 3') (Martínez et al., 2010). PCR reactions were conducted in 30  $\mu$ L: 10  $\mu$ L of ultra pure water, 15  $\mu$ L (0.625 U) of AmpliTaq mix (Applied Biosystems), 0.5  $\mu$ L of each primer (10  $\mu$ M) and 4  $\mu$ L of template DNA (10 ng/ $\mu$ L). The PCR conditions for 16S and CAM markers were: (5 min at 95  $^{\circ}$ C), (30 s at 94  $^{\circ}$ C; 30 s at 46  $^{\circ}$ C; 1 min at 72  $^{\circ}$ C)  $\times$  5, (30 s at 94  $^{\circ}$ C; 30 s at 51  $^{\circ}$ C; 1 min at 72  $^{\circ}$ C)  $\times$  30, and (5 min at 72  $^{\circ}$ C). The PCR conditions for the ITS marker were: (5 min at 95  $^{\circ}$ C), (30 s at 94  $^{\circ}$ C; 30 s at 55  $^{\circ}$ C; 1 min at 72  $^{\circ}$ C)  $\times$  35 and (5 min at 72  $^{\circ}$ C). PCR products were visualized on a 1% agarose gel stained with GelRed Nucleic Acid Stain, 10000X in DMSO (Gentaur, Kampen hout, Belgium). PCR products were sequenced in both directions by Genoscope (CEA Evry, France) and by Genoscreen (Lille, France) on capillary sequencer ABI3730XL.

### 2.4. Microsatellite amplification and genotyping

We used the microsatellite markers developed for *L. breviostris* (Postaire et al., 2015a) and *M. phoenicea* (Postaire et al., 2015b), using the same PCR conditions as Postaire et al. (2015b). Fragment analysis was performed on an ABI Prism 3730 automated sequencer (Applied Biosystems) at the Plateforme Gentyane (INRA, Clermont Ferrand, France). Genotypes were analyzed using GeneMapper v.4 (Applied Biosystems).

### 2.5. Phylogenetic analyses

Sequences were checked and edited using Geneious v.6.0 (Kearse et al., 2012) and deposited in GenBank (Appendix B). Additional 16S, CAM and ITS sequences previously published (Leclère et al., 2007, 2009; Moura et al., 2008, 2012; Postaire et al., 2015c) were retrieved from GenBank (Appendix B). Sequences were aligned using MAFFT (Katoh et al., 2005) and trimmed to the shortest sequence. JModeltest v.2.5 (Darrriba et al., 2012; Posada, 2008) was used to identify the best substitution model based on AICc criterion of each fragment (16S: GTR + I + G; CAM: TPM2uf + I + G; ITS: TIM1 + I + G).

We performed Maximum Likelihood (ML) using PhyML (Geneious plug in; Guindon et al., 2010) and Bayesian inferences (BI) analyses using MrBayes v.3.2 (3 independent runs of: 20  $\times$  10<sup>6</sup> generations, 8 chains, temperature to 0.2, 10% burn in length, sampling every 2  $\times$  10<sup>3</sup> generations; Geneious plug in; Ronquist et al., 2012). Nodes can be considered robust if their posterior probability (PP) is equal or higher to 0.95 for Bayesian reconstruction and when their bootstrap (BS) values are superior to 75% for ML reconstruction (Erixon et al., 2003). These criteria were used hereafter. We used

the same outgroups as Moura et al. (2012) and Postaire et al. (2015c) for 16S tree reconstruction, i.e. the hydrozoans *Schizotricha turqueti* (Billard, 1906) and *S. nana* Peña Cantero et al., 1996.

As assessed by Postaire et al. (2015c), the phylogenetic signals of the three markers are not incongruent and thus we reconstructed two phylogenetic trees: (1) using the whole dataset of 16S unique haplotypes and (2) using a concatenated alignment of the three molecular markers, focusing on *L. breviostris* and *M. phoenicea* (see Section 3).

### 2.6. Delimiting SSHs

All species delimitation methods used in this study have been designed for barcode markers. As the nuclear markers, calmodulin fragment and ITS, have not yet been sequenced for several PSHs of *Aglaophenia*, we only analyzed the 16S marker for all PSHs to delimit secondary species hypotheses (SSHs).

#### 2.6.1. ABGD

We used the ABGD method developed by Puillandre et al. (2012a) on the web server <http://www.wabi.snv.jussieu.fr/public/abgd/abgdweb.html>. After sequence alignment, we computed a matrix of pairwise distances using the K2P model (Kimura, 1980). A graphical representation of the pairwise distance distribution of our dataset showed a narrow barcoding gap between 0.07 and 0.09 divergence. We used  $P_{\min} = 0.001$  and  $P_{\max} = 0.1$  and  $X = 1.0$  as it was the highest value that could be applied.

#### 2.6.2. GMYC models

We used the GMYC method developed by Pons et al. (2006), implemented in R (R Development Core Team, 2008). GMYC analyses require an ultrametric tree (i.e. calibrated with a molecular clock), which was constructed using BEAST v.1.8. (Bouckaert et al., 2014). We used a relaxed log normal clock with a coalescent tree prior as these have been identified as best prior parameters for GMYC analyses (Esselstyn et al., 2012; Monaghan et al., 2009). Monte Carlo Markov chains (MCMC) ran for 35  $\times$  10<sup>6</sup> generations, sampling every 3500 generations. Chains convergence was assessed using Tracer v.1.6 (Rambaut et al., 2014). The consensus tree (maximum clade credibility tree; 10% burn in; tree not presented) was constructed with TreeAnnotator v.1.7 (Rambaut and Drummond, 2013).

To account for uncertainty in species delimitation, we used three applications of the GMYC model: (1) the single threshold species delimitation GMYC model, (2) the multiple threshold species delimitation GMYC model using R (R Development Core Team, 2008) packages *ape* (Paradis et al., 2004) and *splits* (Ezard et al., 2009), and (3) Bayesian GMYC (bGMYC) model developed by Reid and Carstens (2012) in the package *bGMYC*. Using the consensus tree, single and multiple threshold GMYC species delimitation models allow identifying respectively one or several thresholds, dividing coalescent and Yule processes on the tree (Monaghan et al., 2009). As the multiple threshold model allows variation of evolution rates along branches and thus several shift points between Yule and coalescent processes across the phylogenetic tree, we compared the likelihood of the two model outcomes. The Bayesian implementation of the GMYC model (Reid and Carstens, 2012) accounts for uncertainty in the phylogeny and model parameters by sampling trees and conducting MCMC. This application gives marginal probabilities to species hypotheses. As recommended by the authors, we conducted the bGMYC analysis by resampling the tree file generated by BEAST at a lower frequency, resulting in 100 trees (one every 35  $\times$  10<sup>4</sup> generations). Each of them was re run for 5  $\times$  10<sup>4</sup> generations, with 4  $\times$  10<sup>4</sup> generations of burn in and sampling every 100 steps, resulting in 100 new trees per initially sampled tree: in fine, 10<sup>4</sup> new trees were

used. MCMC estimates from each tree were pooled to calculate the probabilities that two leaves in the phylogeny are conspecific. We set the probability of two leaves being conspecific at 0.90 or higher, i.e. each bGMYC cluster has a minimum probability of 0.90 to represent a species. This threshold was used because we observed that these clusters most often corresponded to robust genetic clades and were composed of morphologically rather similar individuals. We excluded lower threshold values as they resulted in outlining larger clades composed of several PSHs with distinct reproductive systems as conspecifics.

### 2.6.3. PTP analyses

We ran a PTP species delimitation analysis in the bPTP web server <http://species.hits.org/ptp/>. As input, we used the maximum likelihood phylogeny of the 16S dataset (see Section 2.5). Outgroups were pruned before conducting the PTP analyses to avoid bias that may arise if some of the outgroup taxa are too distantly related to ingroup taxa (Zhang et al., 2013). We ran the PTP analysis for  $5 \times 10^5$  MCMC generations, with a thinning value of 100, a burn in of 25% and we visually confirmed the convergence of the MCMC chain as recommended by Zhang et al. (2013).

### 2.6.4. Comparison of methods outputs

Since ABGD, GMYC and PTP have a general tendency to over estimate the true number of species from the dataset, we combined their outputs: SSHs were defined considering only the more inclusive clades found by all three methods. These methods make different assumptions and simplifications of evolutionary processes, forcing us to compile their different results into a single evolutionary model that is the basis for the interpretations of SSHs (Carstens et al., 2013). Furthermore, we excluded singletons (i.e. SSHs represented by only one haplotype): we will discuss only SSHs represented by several haplotypes. We chose this conservative approach because we considered that failing to detect potential cryptic species a lesser problematic pitfall than delineating entities that do not represent real evolutionary lineages (Carstens et al., 2013).

### 2.7. Testing SSHs of *L. brevisrostris* and *M. phoenicea* using microsatellite data and nuclear sequence data

For both PSHs, prior to clustering analyses using microsatellite data, repeated multi locus genotypes (MLGs) were identified with GenClone v.2.0 (Arnaud Haond and Belkhir, 2007) using the maximum set of loci for each sampling location. Further analyses were conducted using one representative of each MLG per population.

We used two different methods to assess the genetic clustering of individuals within both PSHs. First, we performed a Discriminant Analysis of Principal Components (DAPC) in the R package *adegenet* (Jombart, 2008; Jombart et al., 2010). DAPC is a non model based method that maximizes the differences between groups while minimizing variation within groups without prior information on individuals' origin. In addition, this method does not assume Hardy Weinberg equilibrium (HWE) or absence of linkage disequilibrium (LD). We used the function `find.clusters()` to assess the optimal number of groups with the Bayesian information criterion (BIC) method (i.e. K with the lowest BIC value should reflect the optimal number of clusters). We tested values of  $K = 1 - 30$ , but BIC values may keep decreasing after the true K value in case of genetic clines and hierarchical structure (Jombart et al., 2010). Furthermore retaining too many discriminant functions with respect to the number of populations can lead to over fitting of data, resulting in spurious discrimination of any set of clusters. Therefore, the rate of decrease in BIC values was visually examined to identify values of K after which BIC values decreased only slightly (Jombart et al., 2010). The `dapc()` function was then executed using the best grouping, retaining axes of PCA sufficient to explain  $\geq 80\%$  of the total

variance. Afterwards, we estimated the number of populations and conducted population assignment of individuals using methods implemented in STRUCTURE v.2.3.4 (Pritchard et al., 2000). The analysis assigns MLGs probabilistically to one of K clusters (K is user defined) so that HWE is maximized and LD minimized in each cluster. For all STRUCTURE analyses, we used the admixture model and the correlated allele frequencies model, without any location or population priors using the following parameters (after validation of chains convergence): three iterations of  $5 \times 10^5$  MCMC generations after an initial burn in of  $5 \times 10^4$  generations for each K, varying from  $K = 1$  to  $K = 14$ . In addition to direct examination of STRUCTURE outputs, Evanno et al. (2005) proposed a method to choose the most likely K by analyzing the second order rate of change of the posterior probability (PP) of the data ( $\Delta K$ ) between successive K values, which was realized using STRUCTURE Harvester v.0.6.94 (Earl and vonHoldt, 2012). Results files were permuted using CLUMPP v.1.1.2 (Jakobsson and Rosenberg, 2007) and visualized using DISTRUCT v.1.1 (Rosenberg, 2004).

#### 2.7.1. *Lytocarpia brevisrostris*

We analyzed the whole dataset (525 MLGs) and calculated pairwise  $F_{ST}$  values between lineages  $\alpha$  and  $\beta$  (see Section 3.3), merging populations in Arlequin v.3.5 (Excoffier et al., 2005). The significance of the observed  $F_{ST}$  statistics was tested using the null distribution generated from  $5 \times 10^3$  non parametric random permutations. We calculated pairwise  $F_{ST}$  values between the clusters identified by STRUCTURE. As sampling was geographically and numerically uneven between lineages, we tested whether microsatellite loci could detect the differentiation between both lineages without an extensive population sampling. Thus, we simulated 10 new datasets composed of all individuals belonging to lineage  $\beta$  previously identified ( $n = 22$ , six from WIO and 16 from SWP) and 22 randomly selected individuals among lineage  $\alpha$ , conserving the same proportions concerning their geographic origin (six from the WIO and 16 from the SWP). For each created dataset, we ran STRUCTURE (see Section 2.7 for parameters) and calculated pairwise  $F_{ST}$  between clusters after convergence.

#### 2.7.2. *Macrorhynchia phoenicea* complex

Prior to microsatellite data analysis, MLGs belonging to under sampled PSHs [i.e. *M. sibogae* (Billard, 1913) and *M. spectabilis* (Allman, 1883)] and locations (i.e. the Scattered Islands except Juan de Nova Island) were pruned from the dataset, as genetic clustering analyses are not powerful enough to handle such small sample sizes. We analyzed the whole dataset (1073 MLGs) and calculated pairwise  $F_{ST}$  values between SSHs and morpho types while merging populations, using Arlequin v.3.5 (Excoffier et al., 2005). The significance of the observed  $F_{ST}$  statistics was tested using the null distribution generated from  $5 \times 10^3$  non parametric random permutations. We calculated pairwise  $F_{ST}$  values between the clusters identified by STRUCTURE.

#### 2.7.3. Comparison to nuclear sequences

After comparison of sequenced based species delimitation and microsatellite data outputs in *L. brevisrostris* and the *M. phoenicea* complex, we observed the congruence of these methods with a phylogenetic reconstruction based on 16S, CAM and ITS concatenated sequences. ML and BI were used (see Section 2.5. for parameters) to reconstruct phylogenetic relationships.

## 3. Results

### 3.1. 16S variability in *Aglaopheniidae* and phylogenetic reconstruction

A total of 37 PSHs were identified in this study. A set of 340 sequences (207 unique haplotypes) of 521 base pairs (bp) was gen

erated and analyzed; this dataset comprised 396 polymorphic sites, 98 identical sites (19.0%), 81.9% pairwise identity and a GC content of 24.6% (base composition: A = 42.5%, C = 12.3%, G = 14.6%, T = 30.6%). Six PSHs were singletons (i.e. PSHs represented by only one haplotype): *Streptocaulus multiseptatus* (Bale, 1915), *S. dolfusii* (Billard, 1924), *Cladocarpus integer* (Sars, 1873), *C. paraformosus* Schuchert, 2000, *A. latecarinata* Allman, 1877, *Lyto carpia* sp.1. The 16S sequences of sampled WIO *M. phoenicea* morpho type A presented a deletion of 95 bp compared to all other Aglaopheniidae sequences, including *M. phoenicea* morpho type B. The reconstructed phylogeny and the results of ABGD, GMYC and PTP analyses are presented in Fig. 2. Before further analyses and after tree reconstructions, all outgroups were pruned from the trees.

As found by Moura et al. (2012) and Postaire et al. (2015c), almost all PSHs were monophyletic, except five: (1) the clade formed by *Gymnangium eximium* (Allman, 1874) and *G. gracilicaule* (Jäderholm, 1903), and (2) the clade formed by *Aglaophenia octodonta* Heller, 1868, *A. pluma* (Linnaeus, 1758) and *A. tubiformis* Marktanner Turneretscher, 1890 (Fig. 2).

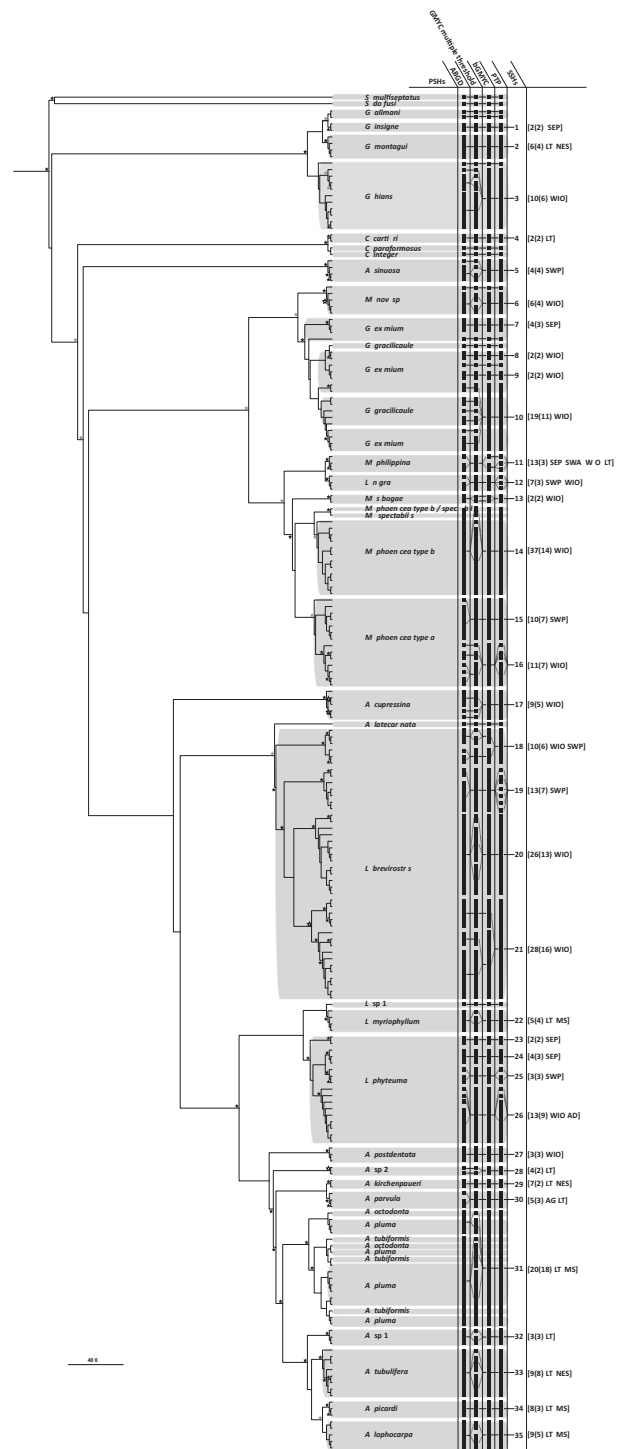
### 3.2. Outputs of species delimitation methods

First, the ABGD method delimited 77 groups (44 clusters and 33 singletons), with a maximum intragroup divergence  $P = 0.010594$  (Fig. 2). Almost all identified groups represented PSHs (48.6%) or clades within PSHs (46%). Two groups of PSHs were merged into ABGD groups: (1) *M. phoenicea* morpho type B with *M. spectabilis*, and (2) *A. pluma*, *A. tubiformis* and *A. octodonta*. Second, the GMYC single threshold model (phylogeny composed of several species with one coalescent time value) was preferred over the null model (single species phylogeny with only coalescent processes;  $P < 0.001$ ). This model identified 15 GMYC clusters (95% Confidence Interval = [6 34]). Based on AICc scores, the multiple threshold model was preferred over the null model ( $P < 0.001$ ) but also performed slightly better than the single threshold model (multiple threshold AICc = 2715; single threshold AICc = 2710); therefore the latter was excluded from further analyses. The multiple threshold model identified four independent switches between speciation and coalescent processes, resulting in 81 entities: 49 GMYC clusters (95% Confidence Interval = [36 51]) and 32 singletons (Fig. 2). Twenty four GMYC clusters (49%) corresponded to robust phylogenetic clades. Eight GMYC clusters (16.3%) corresponded to PSHs whereas all others divided PSHs into several clusters (Fig. 2). Third, the bGMYC analysis identified 53 entities, among which 37 clusters presented a probability  $\geq 0.90$  of being conspecific. Of these clusters, 27 (73%) were phylogenetically robust (Fig. 2). Finally, the PTP analysis identified 63 phylogenetic species, with 33 clusters and 30 singletons.

### 3.3. Identification of SSHs

The comparison of the four species delimitation methods led to the identification of 35 SSHs (Fig. 2, Appendix C). Among these, 19 SSHs (61.3%) matched their corresponding PSHs, while all methods tended to split various PSHs into several lineages. Conversely, some PSHs were merged, highlighting incongruences between morphological and genetic data: *G. eximium*/*G. gracilicaule*, SSH 10; *A. pluma*/*A. octodonta*/*A. tubiformis*, SSH 31; *M. phoenicea* morpho type B/*M. spectabilis*, SSH 14 (Fig. 2, Appendix C), supporting the results of previous phylogenetic studies (Moura et al., 2012 and Postaire et al., 2015c). One PSH (*G. allmani* (Marktanner Turneretscher, 1890)) was divided into two singletons. All other singletons belonged to PSHs also represented by clusters.

Four PSHs [*M. phoenicea* morpho type A, *L. phyteuma* (Kirchenpauer, 1876), *L. breviscostis* and *G. eximium*] were

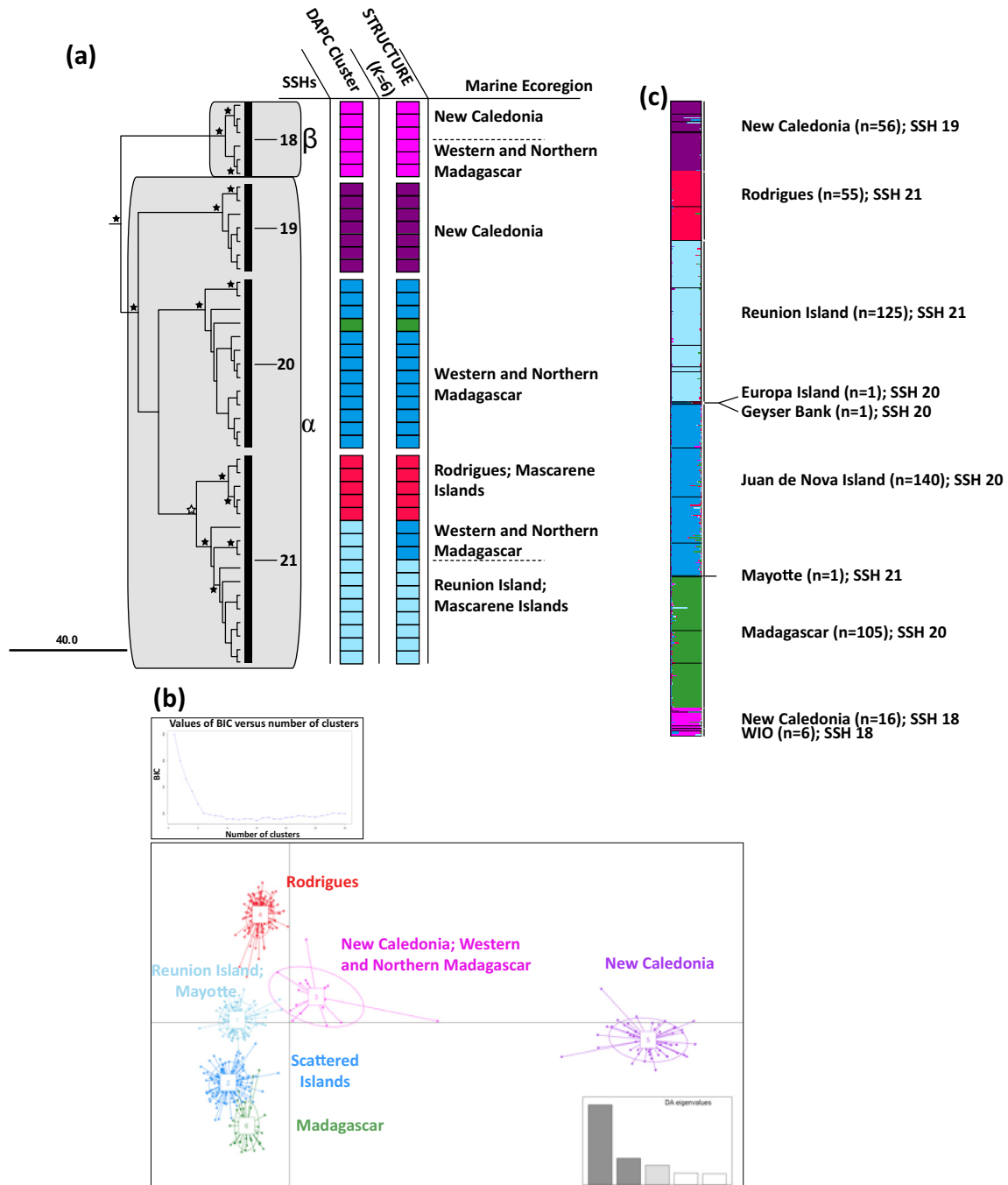


**Fig. 2.** Phylogenetic relationships of Aglaopheniidae species based on 16S sequences. Outgroups were removed. The tree summarizes the results of Maximum Likelihood (ML) and Bayesian Inference (BI) analyses. Black stars indicate tree nodes with a posterior probability (PP)  $\geq 0.95$  and bootstrap support (BS)  $> 75\%$ ; grey stars indicate tree nodes with a PP  $\geq 0.95$ ; blank stars indicate tree nodes with BS  $> 75\%$ . Grey boxes represent primary species hypotheses (PSH; A.: *Aglaophenia*; C.: *Cladocarpus*; G.: *Gymnangium*; L.: *Lytocarpia*; M.: *Macrorhynchia*; S.: *Streptocaulus*). Black vertical bars represent singletons and clusters identified by species delimitation methods (from left to right: the ABGD, the multiple-threshold GMYC, the bGMYC with a 0.90 probability of being conspecific and the PTP analyses). The number juxtaposed to vertical bars is the secondary species hypothesis (SSH) identifier (singletons excluded) with, in brackets, the number of individuals (n), the number of unique haplotypes (Nh) and the sampling provinces (AD: Andaman; AG: Agulhas; AR: Artic; LT: Lusitania; MS: Mediterranean Sea; NES: Northern European seas; SEP: South-Eastern Polynesia; SWA: warm temperate South-Western Atlantic; SWP: South-Western Pacific; WIO: Western Indian Ocean).

composed of several SSHs (Fig. 2, Appendix C). Concerning *L. phyteuma*, these SSHs were partly related to the geographic origin of the samples: two in the SEP (SSHs 23 and 24), one in the SWP (SSH 25) and one in the WIO (SSH 26). This pattern was similar in *M. phoenicea* morpho type A with SSHs 15 and 16 from SWP and WIO, respectively. The PSH *L. brevirostris* was composed of two highly divergent clades that were not entirely related to geography: SSHs 20 and 21 were found in the same geographic region (WIO) and formed a group more related to SSH 19 (SWP; these three SSHs were hereafter referred to as *L. brevirostris* lineage  $\alpha$ ) than to SSH 18 (found both in WIO and SWP;

referred hereafter as *L. brevirostris* lineage  $\beta$ ). Finally, SSHs in *G. eximium* and *G. gracilicaule* were not clearly related neither to geography nor morphology.

Even if the phylogenetic relationships within PSHs *M. phoenicea* *sensu lato* (i.e. *M. phoenicea* morpho types A and B, *M. sibogae* and *M. spectabilis*) were not clear, it is noteworthy that the distinction observed between *M. phoenicea* morpho types seemed relevant. Indeed, two SSHs were found within the *M. phoenicea* morpho type A (SSHs 15 and 16), each corresponding to a biogeographic province, and a third SSH is formed by *M. spectabilis* and *M. phoenicea* morpho type B (SSH 14).



**Fig. 3.** (a) Sub-tree of *Lytocarpia brevirostris* with lineages  $\alpha$  and  $\beta$  and the secondary species hypotheses (SSH) resulting from the different species delimitation methods (black vertical bars) with (b) DAPC and (c) STRUCTURE outputs. Black stars indicate tree nodes with a posterior probability (PP)  $\geq 0.95$  and bootstrap support (BS)  $> 75\%$ ; grey stars indicate tree nodes with a PP  $\geq 0.95$ ; blank stars indicate tree nodes with BS  $> 75\%$ .



**Table 1**

Proportion of missing data (%) in the *L. brevisrostris* microsatellite dataset by locus, sampling province and lineage with the number of multi-locus genotypes in parentheses (n). The 10 loci conserved for the analyses of the whole dataset are in bold.

	Lb01	Lb02	Lb03	Lb04	Lb05	Lb06	Lb07	Lb08	Lb09	Lb10	Lb11	Lb12	Lb14	Lb15	Lb16
WIO $\alpha$ (n = 430)	<b>9.5</b>	<b>6.5</b>	<b>2.8</b>	49.8	<b>6.4</b>	<b>4.2</b>	<b>2.6</b>	<b>5.3</b>	68.1	<b>5.1</b>	<b>6.3</b>	49.3	14.0	11.6	<b>4.4</b>
SWP $\alpha$ (n = 73)	<b>1.4</b>	<b>8.2</b>	<b>0.0</b>	1.4	<b>0.0</b>	<b>1.4</b>	<b>2.7</b>	<b>1.4</b>	1.4	<b>1.4</b>	<b>1.4</b>	1.4	4.1	1.4	<b>1.4</b>
WIO $\beta$ (n = 6)	<b>16.7</b>	<b>0.0</b>	<b>0.0</b>	66.7	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	100.0	<b>0.0</b>	<b>50.0</b>	100.0	83.3	0.0	<b>50.0</b>
SWP $\beta$ (n = 16)	<b>31.3</b>	<b>37.5</b>	<b>0.0</b>	81.3	<b>56.3</b>	<b>0.0</b>	<b>6.3</b>	<b>25.0</b>	100.0	<b>0.0</b>	<b>93.8</b>	81.3	43.8	56.3	<b>25.0</b>
Total (n = 525)	<b>9.1</b>	<b>7.6</b>	<b>2.3</b>	44.2	<b>7.0</b>	<b>3.6</b>	<b>2.7</b>	<b>5.3</b>	60.2	<b>4.4</b>	<b>8.8</b>	44.2	14.3	11.4	<b>5.2</b>

#### 3.4. Comparison of SSHs in *L. brevisrostris* and *M. phoenicea* with microsatellite data

As indicated above, the subtree of *L. brevisrostris* presented two highly divergent lineages  $\alpha$  and  $\beta$ , both composed of individuals sampled in the WIO and the SWP (Fig. 3a). Noteworthy, all loci were amplified but several loci amplified poorly in individuals from SWP: the proportion of missing data by locus according to lineage and origin of individuals is indicated in Table 1 (NB: Lb13 was pruned early from our study due to high scoring incertitude). In fine, eliminating loci with a high proportion of missing data ( $\geq 10\%$ ) resulted in 10 usable loci. High genetic differentiation between lineages  $\alpha$  and  $\beta$  was indicated by differential amplification of loci (Table 1), as well as by the high and significant  $F_{ST}$  value between both lineages calculated using the 10 common loci ( $F_{ST} = 0.16$ ,  $P < 0.001$ ).  $F_{ST}$  values between *L. brevisrostris* SSHs were also high and significant ( $P < 0.001$ ), ranging from 0.15 to 0.24 (Table 2).

DAPC and STRUCTURE analyses revealed that genotypes of *L. brevisrostris* grouped mostly according to their geographic origin. DAPC identified six clusters (Fig. 3b): five corresponding to individuals sampled from different islands or marine ecoregions, while the sixth was composed of individuals from both the WIO and the SWP (corresponding to lineage  $\beta$ ). STRUCTURE analysis also showed that  $\Delta K$  yielded a clear maximum at  $K = 6$ : five clusters corresponded to sampled islands plus a cluster formed by individuals from the WIO and the SWP (corresponding to lineage  $\beta$ ; Fig. 3c).  $F_{ST}$  values between STRUCTURE clusters were all highly significant, ranging from 0.09 ( $P < 0.001$ ) to 0.21 ( $P < 0.001$ ; Table 3). The creation of the 10 artificial datasets, composed evenly by individuals from both lineages and sampling zone, led to the elimination of two additional loci, for the same reasons as explained above. Their STRUCTURE analyses all converged to  $K = 2$ , with mean  $F_{ST} = 0.27^{***}$  (se = 0.0066).

Phylogenetic relationships in the *M. phoenicea sensu lato* subtree were not fully resolved (Fig. 4a), but Bayesian inference indicated that morpho types A and B formed robust clades. Among

**Table 2**

*Lytocarpia brevisrostris* pairwise  $F_{ST}$  values for all pairs of secondary species hypotheses (SSH) with test significance (\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ). Lineage  $\alpha$ : SSHs 19, 20 and 21; lineage  $\beta$ : SSH 18.

	SSH 18	SSH 19	SSH 20
SSH19	0.22***		
SSH 20	0.22***	0.24***	
SSH 21	0.24***	0.24***	0.15***

**Table 3**

*Lytocarpia brevisrostris* pairwise  $F_{ST}$  values for all pairs of STRUCTURE clusters with test significance (\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ).

	Lineage $\beta$ (SSH 18)	Juan de Nova Island	Madagascar	New Caledonia	Rodrigues
Juan de Nova Island	0.10***				
Madagascar	0.16***	0.09***			
New Caledonia	0.11***	0.14***	0.18***		
Rodrigues	0.20***	0.11***	0.19***	0.20***	
Reunion Island	0.16***	0.13***	0.11***	0.18***	0.21***

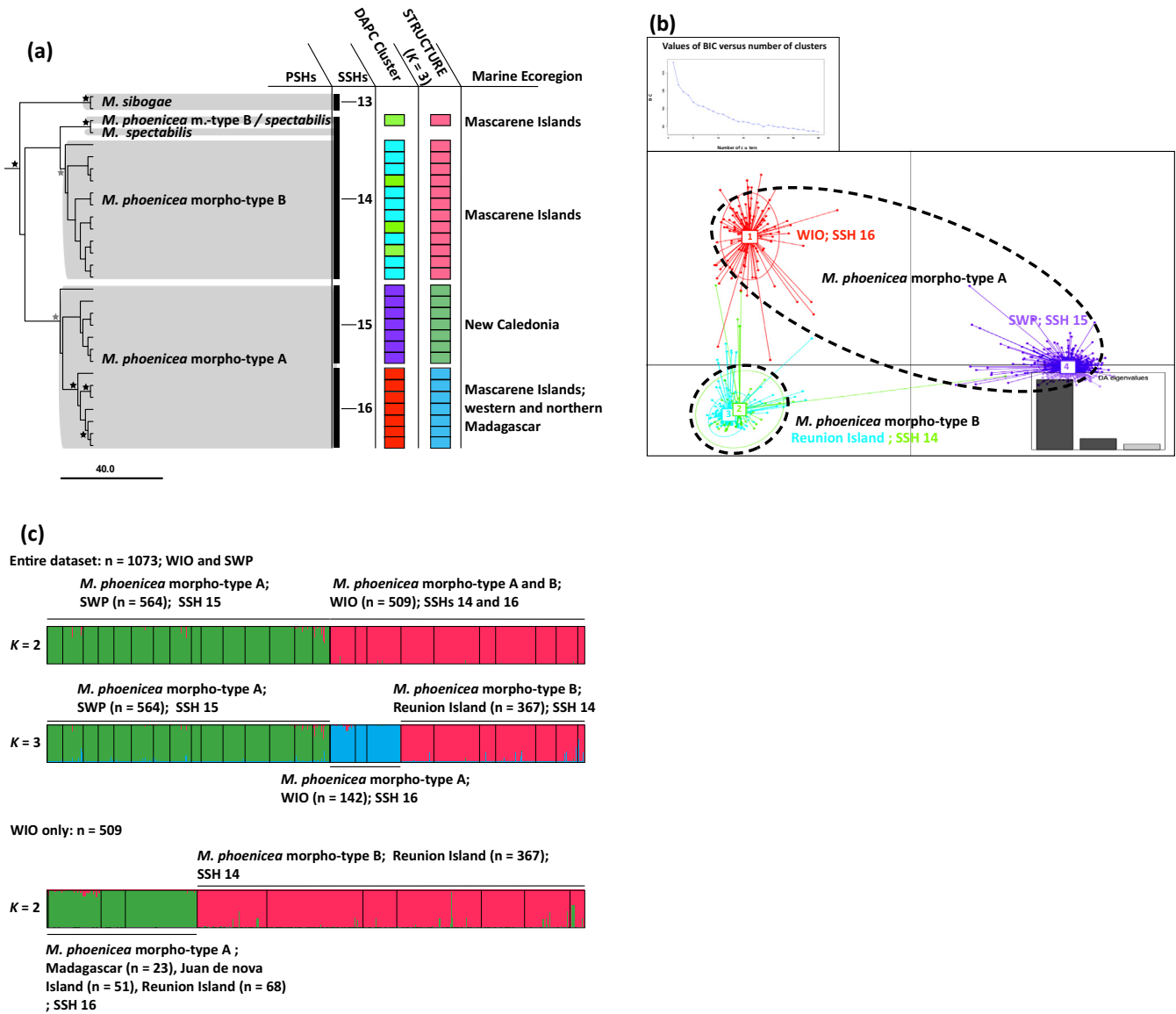
available loci, two failed to amplify for morpho type A and several loci were mono or poly allelic depending on morpho types (Table 4). After eliminating loci with  $\geq 10\%$  of missing data and conserving only those polymorphic in both morpho types, the analysis of the *M. phoenicea* complex was conducted using 8 loci. Pairwise  $F_{ST}$  values between *M. phoenicea* SSHs were extremely high and significant, ranging from 0.27 ( $P < 0.001$ ) to 0.41 ( $P < 0.001$ ; Table 5). Differentiation between morpho types was also strong and significant ( $F_{ST} = 0.33$ ,  $P < 0.001$ ). DAPC analyses did not converge to a single scheme, indicating a possible hierarchical structure in our dataset. However, in the successive number of clusters tested, the initial decline in BIC values slightly slowed at  $K = 4$  (Fig. 4b), but continued decreasing with increasing K. The four clusters corresponded to sampling provinces and morpho types, with morpho type B spread among two clusters (Fig. 4b). This clustering was consistent when increasing the value of K, new clusters being formed by splitting the four original ones. The Bayesian clustering analysis also converged poorly. Indeed, for the entire dataset (n = 1073 MLGs),  $\Delta K$  showed a maximum at  $K = 2$ . The peak corresponded to the clustering of individuals according to their origin: SWP versus WIO (Fig. 4c). At  $K = 3$ , individuals clustered according to SSHs (see Table 5 for  $F_{ST}$  values) and, as this clustering was consistent when increasing K, we decided to use only individuals from the WIO since both morpho types were present in this region. In this subset (n = 509),  $\Delta K$  showed a maximum at  $K = 2$ . At  $K = 2$ , one cluster regrouped all individuals identified as *M. phoenicea* morpho type A (SSH 16), and the second cluster comprised all individuals of *M. phoenicea* morpho type B (SSH 14; Fig. 4c) with a high genetic differentiation ( $F_{ST} = 0.33$ ,  $P < 0.001$ ).

#### 3.5. Comparison of SSHs in *Lytocarpia brevisrostris* and *Macrorhynchia phoenicea* with nuclear sequences

A set of 175 unique haplotypes (Appendix B) of 1870 bp was generated and analyzed (only the subtrees corresponding to *L. brevisrostris* and *M. phoenicea sensu lato* are presented); this dataset comprised 1171 polymorphic sites, 699 identical sites (37.7%), 79.2% pairwise identity and a GC content of 39.5% (base composition: A = 33%, C = 18.6%, G = 20.9%, T = 27.5%).

The subtree containing *L. brevisrostris* sequences presented the same topology and even higher node support than the 16S phylogenetic reconstruction: lineages  $\alpha$  and  $\beta$  were recovered, each containing individuals from both sampling provinces (Fig. 5a).

In contrast, the results concerning *M. phoenicea sensu lato* were not unequivocal (Fig. 5b). On one hand, samples of *M. phoenicea*



**Fig. 4.** (a) Sub-tree of *Macrorhynchia phoenicea* complex with the secondary species hypotheses (SSH) resulting from the different species delimitation methods (black vertical bars) with (b) DAPC and (c) STRUCTURE outputs using first the entire dataset ( $n = 1073$ ) and only individuals from the WIO ( $n = 509$ ) in second analysis. Black stars indicate tree nodes with a posterior probability (PP)  $\geq 0.95$  and bootstrap support (BS)  $> 75\%$ ; grey stars indicate tree nodes with a PP  $\geq 0.95$ ; blank stars indicate tree nodes with BS  $> 75\%$ .

morpho type A formed robust clades according to their origin, corresponding to SSHs 15 and 16. On the other hand, SSH 14, which represented *M. phoenicea* morpho type B and *M. spectabilis*, was polyphyletic when using the concatenated dataset. Indeed, *M. phoenicea* morpho type B and *M. spectabilis* haplotypes clustered in two different partly unsupported clades. Furthermore *M. sibogae* (SSH 13) was no longer external to the clade composed by *M. phoenicea* morpho types A, B and *M. spectabilis* (Fig. 5b). Interestingly, concatenated sequences presented the same structure as microsatellite data: sequences grouped principally according to geographic origin, but with low support.

#### 4. Discussion

We used the mitochondrial marker 16S to study lineage diversity in Aglaopheniidae and detect potential cryptic species. Based on our protocol (i.e. excluding singletons), the combination of several molecular based species delimitation methods identified 35

SSHs among 37 PSHs (Fig. 2, Appendix C). Several SSHs did not correspond to PSHs, revealing potential cases of synonymy and the presence of independent lineages within PSHs. Nevertheless, the majority of PSHs were monophyletic and SSHs were often related to geography (see *L. phyteuma*, *L. brevirostris*, *M. phoenicea* complex, Fig. 2).

##### 4.1. Life cycle and evolutionary lineages

All methods used split several PSHs into independent lineages. These lineages might reflect deep divergences occurring between populations across PSHs distributions. Indeed, high diversity of independent lineages in widely distributed hydrozoan species has been documented in *Plumularia setacea* (Linnaeus, 1758), another brooding hydrozoan (Schuchert, 2014), as in other hydrozoans (e.g. Folino Rorem et al., 2009; Moura et al., 2008). In *Plumularia setacea*, each of the 10 sampled regions presented a highly divergent, geographically delimited lineage lacking clear morphological distinctions (Schuchert, 2014). As individuals were sampled

**Table 4**

Allelic range per microsatellite locus for both *M. phoenicea* morpho-types with the number of multi-locus genotypes in parentheses (n). The eight loci conserved for the analyses of the whole dataset are in bold.

	Morpho-type A (n = 706)	Morpho-type B (n = 367)
Mp01	No amplification	119–164
Mp02	No amplification	106–123
Mp03	<b>153–194</b>	<b>153–165</b>
Mp04	<b>150–189</b>	<b>155–171</b>
Mp05	<b>228–256</b>	<b>230–256</b>
Mp06	<b>268–324</b>	<b>308–333</b>
Mp07	143–162	140–164
Mp08	<b>155–167</b>	<b>155–167</b>
Mp09	<b>176–196</b>	<b>186–194</b>
Mp10	<b>241–347</b>	<b>236–324</b>
Mp11	301	289–324
Mp12	136	135–171
Mp13	154	138–164
Mp14	182	177–188
Mp15	202	202–225
Mp16	212	197–218
Mp17	114–159	132
Mp18	127–180	144
Mp19	153–174	159
Mp20	192–204	201
Mp21	271–305	284
Mp22	119–129	121
Mp23	<b>126–168</b>	<b>136–154</b>
Mp24	139–202	143
Mp25	157–169	166
Mp26	216–242	233

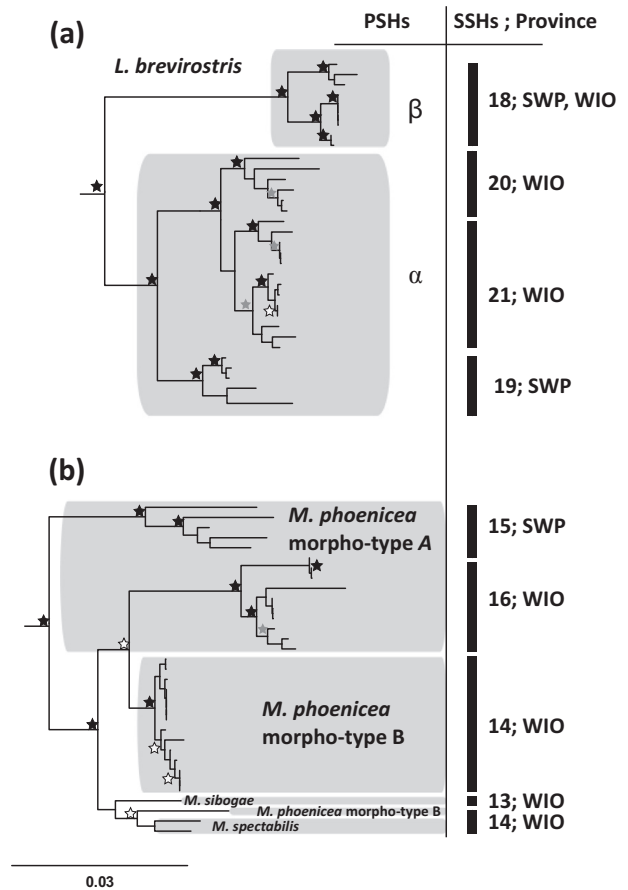
**Table 5**

*Macrorhynchia phoenicea* complex pairwise  $F_{ST}$  values for pairs of SSHs with test significance (\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ).

	SSH 14	SSH 15
SSH 15	0.41***	
SSH 16	0.34***	0.27***

in similar ecosystems, the author argued that this pattern was likely the outcome of a low dispersal capacity of this taxon rather than the result of morphological convergence related to ecological constraints. The majority of Aglaopheniidae morpho species are brooders: larvae are released only when mature and ready to settle (Boero and Bouillon, 1993; Bouillon et al., 2006). Incubating larvae is thought to limit gene flow, but the relationship between pelagic larval duration and dispersal capacity measured via population genetics is not straightforward (Ayre and Hughes, 2000; Fairby and Barber, 2012; Shanks, 2009; Siegel et al., 2008; Teske et al., 2007; Weersing and Toonen, 2009). Indeed, even when considering species with potentially high dispersal capacities, genetic exchanges across ocean basins can still be constrained due to the presence of biogeographic barriers (Ayre et al., 2009; Muths et al., 2011; Ridgway and Sampayo, 2005). However, in ophiuroids, low phylogenetic diversity across ocean basins was associated with the presence of a planktotrophic larva, whereas taxa with lecithotrophic larvae presented important cryptic diversity and allopatric speciation over relatively small geographic scales (Hoareau et al., 2013). Thus, potentially low dispersal capacity appears to facilitate vicariance and allopatric speciation in marine organisms (Paulay and Meyer, 2002).

In the present study, *G. hians* (Busk, 1852) and *M. philippina* Kirchenpauer, 1872 are not brooders but present a reduced medusa stage called medusoid (Bourmaud and Gravier Bonnet, 2004). These two morpho species were represented by few SSHs, composed of individuals from distant biogeographical provinces. In contrast, brooding morpho species with hypothetically lower dispersal capacities, sampled over large geographic scales,



**Fig. 5.** Subtrees presenting phylogenetic relationships within (a) *Lytocarpia brevisrostris* and (b) *Macrorhynchia phoenicea* complex based on the concatenated dataset of 16S, calmodulin and ITS sequences (left). Primary species hypotheses (PSH; grey boxes) are represented and corresponding secondary species hypotheses (SSH) delimited using the 16S only.

presented higher numbers of SSHs: *L. phyteuma* and *L. brevisrostris* were composed of several robust clades, some occurring within the same biogeographic provinces or even sympatrically at sampling sites. These results suggest that larval brooding enhances diversification opportunities in Aglaopheniidae by limiting a taxon's dispersal capacities and its populations' connectivity. While the majority of Aglaopheniidae morpho species are thought to present global distributions (Costello et al., 2013), the extensive lineage diversity uncovered within these taxa (Moura et al., 2012, this study) indicates that they are more likely to represent mosaics of independent meta populations, restricted to small geographic areas, forming cryptic species similarly to some other cnidarians, gastropods and understudied cosmopolitan taxa (Aurelle et al., 2011; Jörger et al., 2012; Mokhtar Jamaï et al., 2011; Payo et al., 2013).

#### 4.2. Sequence based species delimitation methods

The species delimitation methods used here reveal synonymies of several PSHs: *A. pluma*, *A. tubiformis* and *A. octodonta* do not form monotypic clusters but are regrouped into a single SSH. Unfortunately, only one marker was available for these PSHs, and their taxonomic status thus remains unresolved. However, their synonymy was already predicted by Moura et al. (2012) and is supported by their strong morphological similarity and their low genetic diversity. *Gymnangium eximium* and *G. gracilicaule* present another example of PSHs with ambiguous taxonomic status, which is

currently under revision using both morphological and molecular data (Ronowicz et al., submitted for publication).

Previous phylogenetic studies of other cnidarian groups have revealed complex species relationships without clear boundaries, possible hybridization and high cryptic diversity (e.g. Reimer et al., 2007; Souter, 2010), highlighting the difficulty of using morphological characters to delimit species in this phylum. Likewise, hydrozoans, or at least Aglaopheniidae, present a good case where an integrative approach is necessary to settle species limits by combining as much information as possible (Padial et al., 2010; Schlick Steiner et al., 2010) to avoid overlooking cryptic diversity [see an example for cave spiders in Hedin (2015)]. Noteworthy, even if some morphological characters do not represent phylogenetic relationships in this family, most PSHs do form robust monophyletic clades (Moura et al., 2012; Postaire et al., 2015c); therefore morphological characters are valuable in the identification of these taxonomic entities, even if they may encompass more than a single species.

However, before considering each identified SSHs as a species, the priors and specificities of each delimitation method must be considered. First of all, ABGD and GMYC methods might fail to recognize clades undergoing rapid radiation or recently formed species (Puillandre et al., 2012a; Reid and Carstens, 2012). Furthermore, GMYC models assume that the studied taxa do not present a meta population structure (i.e. each identified cluster of haplotypes is a meta population in itself), while ABGD is only based on a measure of distance between sequences, without considering phylogenetic relationships. Also, both methods do not ponder possible introgressions. As it is very unlikely that Aglaopheniidae PSHs are represented by a single evolutionary lineage across multiple ocean basins, SSHs that correspond to the sampling regions might reflect complex population genetic structuring and incomplete lineage sorting (Talavera et al., 2013). Furthermore, all the methods used are designed for barcoding markers, i.e. 16S in this study. Some other universal markers exist for hydrozoans, but they are either too conserved for these methods (supposed absence of a barcoding gap; Baba et al., 1984; Lindner et al., 2011) or present too high intra specific variation (Coleman, 2003; Merino Serrais et al., 2012) to be useful as barcodes. In addition, these methods require samples of the entire species range, with several individuals from each sampling site (Lim et al., 2011; Powell, 2012; Talavera et al., 2013). For most studied PSHs, however, distribution ranges are roughly estimated or inaccurate. As sequence based species delimitation methods identified several entities within PSHs, finer morphological measures, ecological and population genetics data are required. Indeed, population genetics is particularly helpful in determining whether lineages are interbreeding and thus represent biological species (Mayr, 1942).

In our case, a conservative approach would be to consider as SSHs each robust clade that corresponds to a major biogeographic province, or even ecoregion, and complement species delineation methods with population genetics studies and other data types. We covered several of these aspects for two PSHs: *L. brevirostris* and *M. phoenicea sensu lato*.

#### 4.3. Integrative taxonomy in Aglaopheniidae

Our study reveals the extent of the hidden diversity in this family. In the case of *L. brevirostris*, the use of microsatellite data confirmed (1) the identification of several SSHs, while merging some of them, and (2) the existence of sympatric highly divergent lineages within this morpho species (lineages  $\alpha$  and  $\beta$ ), with little gene flow between them. However, when used in species with low gene flow and high population differentiation, the differentiation levels between species and between isolated populations

within species may be indistinguishable. Nonetheless, the phylogenetic relationships reconstructed with nuclear markers supported the results of both mitochondrial and microsatellite data. Thus, considering the congruent outcomes of (1) species delimitation methods, (2) microsatellite data, (3) nuclear sequence markers and (4) the fact that individuals belonging to different SSHs occurred sympatrically, we conclude that this PSH is at least composed of two true biological cryptic species: *L. brevirostris*  $\alpha$  and *L. brevirostris*  $\beta$ .

Using microsatellite data was also conclusive for the *M. phoenicea* complex but needs further exploration: population genetics requires extensive geographic and taxonomic sampling, which we could not achieve for *M. sibogae* and *M. spectabilis*. The status of the former PSH is quite variable depending on authors: *M. sibogae* has been considered either a valid PSH (Billard, 1913; Gravier Bonnet and Fontaine, 1981) or a synonym of *M. phoenicea* (Di Camillo et al., 2009). The incongruence between mitochondrial and nuclear data does not help to clarify this issue. It is worth noting, however, that *M. phoenicea* morpho type B and *M. spectabilis* share a 16S haplotype, supporting their potential synonymy or at least high relatedness. Microsatellite data were useful to clarify the relationships between both *M. phoenicea* morpho types. Even if the clustering scheme and nuclear sequences reflected the mixing of phylogenetic (i.e. differentiation between SSHs and morpho types) and population genetics information (i.e. differentiation between SWP and WIO), all available information confirmed the differentiation between morpho types A and B. Schuchert (2003) already suspected two variants of *M. phoenicea* to be two full species in the Kei Islands (Maluku Province, Indonesia). Thus, as (1) morpho types A and B occur sympatrically in the WIO, (2) they present a reciprocal monophyly, (3) microsatellite loci present discrepancies in amplification and allelic richness, and (4) genetic differentiation is extremely high in sympatric populations, we conclude that *M. phoenicea sensu stricto* is composed of at least two biological species with probable distinct microhabitats. These two species deserve the attention of taxonomists; morpho type A from SWP might either represents another true species or a highly divergent lineage of *M. phoenicea* morpho type A. Our field observations are congruent with genetic data, showing the importance of gathering ecological information (even succinct) when studying taxonomy.

## 5. Conclusions

In this study we employed species delimitation tools to assess the congruence between genetic data and morphological classification in hydrozoans from the Aglaopheniidae family. We identified several independent lineages in nominal morpho species, representing true cryptic species. These results, based on the congruence of sequence and microsatellite data, reveal part of the hidden diversity within the Aglaopheniidae family. Our results underline the relevance of integrative taxonomy in Aglaopheniidae and have direct implications for taxonomy and conservation, highlighting that the current Aglaopheniidae register might seriously underestimate the true diversity within this family.

## Acknowledgements

This work was supported by the Laboratoire d'Excellence CORAIL. Hydrozoan sampling in New Caledonia (HM) was carried out during Cabelo (doi: <http://dx.doi.org/10.17600/14003700>) and Bibelot (doi: <http://dx.doi.org/10.17600/13100100>) oceanographic campaigns on board of RV Alis (IRD). Sampling in Reunion Island (HM, BP, HB, CAFB) was supported by program HYDROSOOI (Labex CORAIL fund); in Madagascar (HM) supported by project

Biodiversity (POCT FEDER fund) and in Juan de Nova (HM) by program Biorecie (POCT FEDER fund). We thank Lucie Penin for samples from French Polynesia and acknowledge the Plateforme Gentyane of the Institut National de la Recherche Agronomique (INRA, Clermont Ferrand, France) for guidance and technical support. The first author was financially supported by a PhD contract from the Sciences Technologie Santé doctoral school of the Université de La Réunion.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ympcv.2016.08.013>.

## References

- Addamo, A.M., Reimer, J.D., Taviani, M., Freiwald, A., Machordom, A., 2012. *Desmophyllum dianthus* (Esper, 1794) in the scleractinian phylogeny and its intraspecific diversity. *PLoS ONE* 7, e50215.
- Agapow, P.M., Bininda-Emonds, O.R.P., Crandall, K.A., Gittleman, J.L., Mace, G.M., Marshall, J.C., Purvis, A., 2004. The impact of species concept on biodiversity studies. *Quart. Rev. Biol.* 79, 161–179.
- Arnaud-Haond, S., Belkhir, K., 2007. GENCLONE: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Mol. Ecol. Notes* 7, 15–17.
- Aurelle, D., Ledoux, J.B., Rocher, C., Borsari, P., Chenuil, A., 2011. Phylogeography of the red coral (*Corallium rubrum*): inferences on the evolutionary history of a temperate gorgonian. *Genetica* 139, 855–869.
- Ayre, D.J., Minchinton, T.E., Perrin, C., 2009. Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Mol. Ecol.* 18, 1887–1903.
- Ayre, D.J., Hughes, T.P., 2000. Genotypic diversity and gene flow in brooding and spawning corals along the Great Barrier Reef, Australia. *Evolution* 54, 1590–1605.
- Baba, M.L., Goodman, M., Berger-Cohn, J., Demaille, J.G., Matsuda, G., 1984. The early adaptive evolution of calmodulin. *Mol. Biol. Evol.* 1, 442–455.
- Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, Ben., Ferrer, E.A., 2011. Has the Earth's sixth mass extinction already arrived? *Nature* 471, 51–57.
- Bavestrello, G., Puce, S., Cerrano, C., Sarà, M., 2000. Phenotypic plasticity in hydrozoans: morph reversibility. *Riv. Biol.* 93, 283–294.
- Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K., Das, I., 2007. Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* 22, 148–155.
- Billard, A., 1913. Les hydroïdes de l'expédition du Siboga. I. Plumulariidae. *Rés. Explor. Siboga 1913 Monogr.* 7, 1–115.
- Boero, F., Bouillon, J., 1993. Zoogeography and life cycle patterns of Mediterranean hydromedusae (Cnidaria). *Biol. J. Linn. Soc.* 48, 239–266.
- Boissin, E., Féral, J.-P., Chenuil, A., 2008. Defining reproductively isolated units in a cryptic and syntopic species complex using mitochondrial and nuclear markers: the brooding brittle star, *Amphipholis squamata* (Ophiuroidea). *Mol. Ecol.* 17, 1732–1744.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10, e1003537.
- Bouillon, J., Gravili, C., Pagès, F., Gili, J.-M., Boero, F., 2006. An Introduction to Hydrozoa. *Mémoires du Muséum d'Histoire Naturelle*, Paris, 591 pp.
- Bourmaud, C.A.-F., Gravier-Bonnet, N., 2004. Medusoid release and spawning of *Macrorhynchia philippina* Kirchenpauer, 1872 (Cnidaria, Hydrozoa, Aglaopheniidae). *Hydrobiologia* 530 (531), 365–372.
- Briggs, J.C., 2005. Coral reefs: conserving the evolutionary sources. *Biol. Conserv.* 126, 297–305.
- Carstens, B.C., Pelletier, T.A., Reid, N.M., Satler, J.D., 2013. How to fail at species delimitation. *Mol. Ecol.* 22, 4369–4383.
- Castelin, M., Lambourdière, J., 2010. Hidden diversity and endemism on seamounts: focus on poorly dispersive neogastropods. *Biol. J. Linn. Soc.* 100, 420–438.
- Coleman, A.W., 2003. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends Genet.* 19, 370–375.
- Costello, M.J., Bouchet, P., Boxshall, G., Fauchald, K., Gordon, D., Hoeksema, B.W., Poore, G.C.B., van Soest, R.W.M., Stöhr, S., Walter, T.C., Vanhoorne, B., Decock, W., Appeltans, W., 2013. Global coordination and standardisation in marine biodiversity through the World Register of Marine Species (WoRMS) and related databases. *PLoS ONE* 8, e51629.
- Cunningham, C.W., Buss, L.W., 1993. Molecular evidence for multiple episodes of paedomorphosis in the family Hydractiniidae. *Biochem. Syst. Ecol.* 21, 57–69.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9, 772.
- Daugherty, C.H., Cree, A., Hay, J.M., Thompson, M.B., 1990. Neglected taxonomy and continuing extinctions of tuatara (*Sphenodon*). *Nature* 347, 177–178.
- Dayrat, B., 2005. Towards integrative taxonomy. *Biol. J. Linn. Soc.* 85, 407–415.
- De Queiroz, K., 2007. Species concepts and species delimitation. *Syst. Biol.* 56, 879–886.
- De Queiroz, K., 2005. Ernst Mayr and the modern concept of species. *Proc. Natl. Acad. Sci. USA* 102, 6600–6607.
- De Queiroz, K., 1992. Phylogenetic definitions and taxonomic philosophy. *Biol. Philos.* 7, 295–313.
- Dellicour, S., Flot, J.-F., 2015. Delimiting species-poor data sets using single molecular markers: a study of barcode gaps, Haplowebs and GMYC. *Syst. Biol.* 64 (6), 900–908.
- Di Camillo, C.G., Puce, S., Bavestrello, G., 2009. *Macrorhynchia* species (Cnidaria: Hydrozoa) from the Bunaken Marine Park (North Sulawesi, Indonesia) with a description of two new species. *Ital. J. Zool.* 76 (2), 208–228. <http://dx.doi.org/10.1080/11250000802182018>.
- Duda Jr, T.F., Bolin, M.B., Meyer, C.P., Kohn, A.J., 2008. Hidden diversity in a hyperdiverse gastropod genus: discovery of previously unidentified members of a *Conus* species complex. *Mol. Phylogenet. Evol.* 49, 177–186.
- Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361.
- Erixon, P., Svennblad, B., Britton, T., Oxelman, B., 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Syst. Biol.* 52, 665–673.
- Esselstyn, J.A., Evans, B.J., Sedlock, J.L., Anwarali Khan, F.A., Heaney, L.R., 2012. Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proc. R. Soc. B* 279 (1743), 3678–3686. <http://dx.doi.org/10.1098/rspb.2012.0705>.
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14, 2611–2620.
- Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform.* 1, 47–50 (Online).
- Ezard, T., Fujisawa, T., Barraclough, T.G., 2009. Splits: Species' Limits by Threshold Statistics. R package version 1.0-11/r29.
- Faurby, S., Barber, P.H., 2012. Theoretical limits to the correlation between pelagic larval duration and population genetic structure. *Mol. Ecol.* 21, 3419–3432.
- Folino-Rorem, N.C., Darling, J.A., D'Ausilio, C.A., 2009. Genetic analysis reveals multiple cryptic invasive species of the hydrozoan genus *Cordylophora*. *Biol. Invas.* 11 (8), 1869–1882.
- Fontaneto, D., Iakovenko, N., Eyres, I., Kaya, M., Wyman, M., Barraclough, T.G., 2010. Cryptic diversity in the genus *Adineta* Hudson & Gosse, 1886 (Rotifera: Bdelloidea: Adinetidae): a DNA taxonomy approach. *Hydrobiologia* 662, 27–33.
- Fujita, M.K., Leaché, A.D., Burbrink, F.T., McGuire, J.A., Moritz, C., 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol. Evol.* 27, 480–488.
- Govindarajan, A.F., Halanych, K.M., Cunningham, C.W., 2005. Mitochondrial evolution and phylogeography in the hydrozoan *Obelia geniculata* (Cnidaria). *Mar. Biol.* 146 (2), 213–222.
- Gravier-Bonnet, N., Bourmaud, C.A.-F., 2006a. Hydroïds (Cnidaria, Hydrozoa) of coral reefs: preliminary results on community structure, species distribution and reproductive biology in Juan de Nova Island (Southwest Indian Ocean). *WIO J. Mar. Sci.* 5, 123–132.
- Gravier-Bonnet, N., Bourmaud, C.A.-F., 2006b. Hydroïds (Cnidaria, Hydrozoa) of coral reefs: preliminary results on community structure, species distribution and reproductive biology in the île Glorieuses (Southwest Indian Ocean). In: *Proceedings of 10th International Coral Reef Symposium*, pp. 188–196.
- Gravier-Bonnet, N., Bourmaud, C.A.-F., 2012. Hydroïds (Cnidaria, Hydrozoa) of Baa atoll (Indian Ocean, Maldives Archipelago). *Atoll Res. Bull.* 590, 85–123.
- Gravier-Bonnet, N., Fontaine, M.F., 1981. *Lytocarpus sibogae* Billard, 1913, espèce nouvelle pour l'océan Indien (Coelenterata, Hydrozoa, Plumulariidae). *Téthys* 10, 103–116.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321.
- Hamilton, C.A., Hendrixson, B.E., Brewer, M.S., Bond, J.E., 2014. An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: a case study of the North American tarantula genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). *Mol. Phylogenet. Evol.* 71, 79–93.
- Hausdorf, B., Wilkens, H., Strecker, U., 2011. Population genetic patterns revealed by microsatellite data challenge the mitochondrial DNA based taxonomy of *Astyanax* in Mexico (Characidae, Teleostei). *Mol. Phylogenet. Evol.* 60 (1), 89–97.
- Hausdorf, B., Hennig, C., 2010. Species delimitation using dominant and codominant multilocus markers. *Syst. Biol.* 59, 491–503.
- Heath, T.A., Hedtke, S.M., Hillis, D.M., 2008. Taxon sampling and the accuracy of phylogenetic analyses. *J. Syst. Evol.* 46 (3), 239–257.
- Hebert, P., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci. USA* 101 (41), 14812–14817.
- Hebert, P.D.N., Ratnasingham, S., de Waard, J.R., 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Philos. Trans. Roy. Soc., B* 270, S96–S99.
- Hedin, M., 2015. High-stakes species delimitation in eyeless cave spiders (Cicurina, Dictynidae, Araneae) from central Texas. *Mol. Ecol.* 24 (2), 346–361. <http://dx.doi.org/10.1111/mec.13036>.

- Hoareau, T.B., Boissin, E., Paulay, G., Bruggemann, J.H., 2013. The Southwestern Indian Ocean as a potential marine evolutionary hotspot: perspectives from comparative phylogeography of reef brittle-stars. *J. Biogeogr.* 40, 2167–2179.
- Hoelzer, G.A., 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees revisited. *Evolution* 51, 622.
- Huelsken, T., Keyse, J., Liggins, L., Penny, S., Tremblay, E.A., Riginos, C., 2013. A novel widespread cryptic species and phylogeographic patterns within several giant clam species (Cardiidae: *Tridacna*) from the Indo-Pacific Ocean. *PLoS ONE* 8, e80858.
- Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806.
- Jombart, T., 2008. ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405.
- Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11, 94.
- Jörger, K.M., Norenburg, J.L., Wilson, N.G., Schrödl, M., 2012. Barcoding against a paradox? Combined molecular species delineations reveal multiple cryptic lineages in elusive meiofaunal sea slugs. *BMC Evol. Biol.* 12, 245.
- Katoh, K., Kuma, K.-I., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucl. Acid Res.* 33, 511–518.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A.J., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649.
- Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120.
- Kingman, J.F.C., 1982. On the genealogy of large populations. *J. Appl. Probab.* 19, 27.
- Knowlton, N., 1993. Sibling species in the sea. *Annu. Rev. Ecol. Syst.* 24, 189–216.
- Ladner, J.T., Palumbi, S.R., 2012. Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes. *Mol. Ecol.* 21, 2224–2238.
- Leclère, L., Schuchert, P., Manuel, M., 2007. Phylogeny of the Plumularioidea (Hydrozoa, Leptothecata): evolution of colonial organisation and life cycle. *Zool. Scripta.* 36 (4), 371–394.
- Leclère, L., Schuchert, P., Cruaud, C., Couloux, A., Manuel, M., 2009. Molecular phylogenetics of Thecata (Hydrozoa, Cnidaria) reveals long-term maintenance of life history traits despite high frequency of recent character changes. *Syst. Biol.* 58 (5), 509–526.
- Lim, G.S., Balke, M., Meier, R., 2011. Determining species boundaries in a world full of rarity: singletons, species delimitation methods. *Syst. Biol.* 61, 165–169.
- Lindner, A., Cairns, S.D., Cunningham, C.W., 2008. From offshore to onshore: multiple origins of shallow-water corals from deep-sea ancestors. *PLoS ONE* 3, e2429.
- Lindner, A., Govindarajan, A.F., Migotto, A.E., 2011. Cryptic species, life cycles, and the phylogeny of *Clytia* (Cnidaria: Hydrozoa: Campanulariidae). *Zootaxa* 2980, 23–36.
- Lohse, K., 2009. Can mtDNA barcodes be used to delimit species? A response to Pons et al. (2006). *Syst. Biol.* 58, 439–442.
- Margules, C.R., Pressey, R.L., 2000. Systematic conservation planning. *Nature* 405 (6783), 243–253.
- Martinez, D.E., Iñiguez, A.R., Percell, K.M., Willner, J.B., Signorovitch, J., Campbell, R. D., 2010. Phylogeny and biogeography of *Hydra* (Cnidaria: Hydrozoa) using mitochondrial and nuclear DNA sequences. *Mol. Phylogenet. Evol.* 57, 403–410.
- Mayden, R.L., 2002. On biological species, species concepts and individuality in the natural world. *Fish Fish.* 3, 171–196.
- Mayr, E., 1942. *Systematics and the Origin of Species, From the Viewpoint of a Zoologist.* Harvard University Press.
- McFadden, C.S., Brown, A.S., Brayton, C., Hunt, C.B., van Ofwegen, L.P., 2014. Application of DNA barcoding in biodiversity studies of shallow-water octocorals: molecular proxies agree with morphological estimates of species richness in Palau. *Coral Reefs* 33, 275–286.
- Merino-Serrais, P., Casado-Amezúa, P., Ocaña, Ó., Templado, J., Machordom, A., 2012. Slight genetic differentiation between western and eastern limits of *Astroides calycularis* (Pallas, 1776) (Anthozoa, Scleractinia, Dendrophylliidae) distribution inferred from COI and ITS sequences. *Graellsia* 68, 207–218.
- Meroz-Fine, E., Brickner, I., Loya, Y., Ilan, M., 2003. The hydrozoan coral *Millepora dichotoma*: speciation or phenotypic plasticity? *Mar. Biol.* 143, 1175–1183.
- Miglietta, M.P., Piraino, S., Kubota, S., Schuchert, P., 2007. Species in the genus *Turritopsis* (Cnidaria, Hydrozoa): a molecular evaluation. *J. Zool. System* 45, 11–19.
- Miglietta, M.P., Schuchert, P., Cunningham, C.W., 2009. Reconciling genealogical and morphological species in a worldwide study of the Family Hydractiniidae (Cnidaria, Hydrozoa). *Zool. Scripta.* 38, 403–430.
- Mokhtar-Jamaï, K., Pascual, M., Ledoux, J.B., Coma, R., Féral, J.-P., Garrabou, J., Aurelle, D., 2011. From global to local genetic structuring in the red gorgonian *Paramuricea clavata*: the interplay between oceanographic conditions and limited larval dispersal. *Mol. Ecol.* 20, 3291–3305.
- Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D.J.G., Lees, D.C., Ranaivosolo, R., Eggleton, P., Barraclough, T.G., Vogler, A.P., 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Syst. Biol.* 58, 298–311.
- Moritz, C., 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. *Syst. Biol.* 51, 238–254.
- Moritz, C., 1999. Conservation units and translocations: strategies for conserving evolutionary processes. *Heredity* 130, 217–228.
- Moura, C.J., Cunha, M.R., Porteiro, F.M., Rogers, A.D., 2012. A molecular phylogenetic appraisal of the systematics of the Aglaopheniidae (Cnidaria: Hydrozoa, Leptothecata) from the north-east Atlantic and west Mediterranean. *Zool. J. Linn. Soc.* 164, 717–727.
- Moura, C.J., Harris, D.J., Cunha, M.R., Rogers, A.D., 2008. DNA barcoding reveals cryptic diversity in marine hydroids (Cnidaria, Hydrozoa) from coastal and deep-sea environments. *Zool. Scripta.* 37 (1), 93–108.
- Muths, D., Tessier, E., Gouws, G., Craig, M.T., Mwale, M., Mwaluma, J., Mwandya, A., Bourjea, J., 2011. Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the SW Indian Ocean. *Mar. Ecol. Prog. Ser.* 443, 167–180.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A., Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858.
- Niemiller, M.L., Near, T.J., Fitzpatrick, B.M., 2011. Delimiting species using multilocus data: diagnosing cryptic diversity in the southern cavefish, *Typhlichthys subterraneus* (Teleostei: Amblyopsidae). *Evolution* 66, 846–866.
- Padial, J.M., Miralles, A., la Riva De, I., Vences, M., 2010. The integrative future of taxonomy. *Front. Zool.* 7, 16.
- Palumbi, S.R., 1994. Genetic divergence, reproductive isolation, and marine speciation. *Annu. Rev. Ecol. Syst.* 25, 547–572.
- Paradis, E., Clavier, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20, 289–290.
- Paulay, G., Meyer, C.P., 2002. Diversification in the tropical Pacific: comparisons between marine and terrestrial systems and the importance of founder speciation. *Integr. Comp. Biol.* 42, 922–934.
- Payo, D.A., Leliert, F., Verbruggen, H., D'hondt, S., Calumpong, H.P., De Clerck, O., 2013. Extensive cryptic species diversity and fine-scale endemism in the marine red alga *Portieria* in the Philippines. *Philos. Trans. Roy. Soc., B* 280 (1753), 20122660.
- Peña Cantero, A.L., Vervoort, W., Svoboda, A., 1996. Species of *Schizotricha* Allman, 1883 (Cnidaria, Hydrozoa) from recent antarctic expeditions with RV Tolarstern, with the description of a new species. *Zool. Meded.* 70 (28), 411–435.
- Pfenninger, M., Schwenk, K., 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evol. Biol.* 7, 6.
- Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Suhlín, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55, 595–609.
- Posada, D., 2008. JModelTest: Phylogenetic model averaging. *Mol. Biol. Evol.* 25, 1253–1256.
- Postaire, B., Aurelle, D., Bourmaud, C.A.-F., Bruggemann, J.H., Magalon, H., 2015a. Isolation and characterisation of 16 microsatellite loci from a widespread tropical hydrozoan, *Lytocarpia brevirostris* (Busk, 1852). *Conserv. Genet. Resour.* 7, 505–507.
- Postaire, B., Aurelle, D., Bourmaud, C.A.-F., Bruggemann, J.H., Magalon, H., 2015b. Isolation and characterisation of 26 microsatellite loci from a widespread tropical hydrozoan, *Macrorhynchia phoenicea* (Leptothecata, Aglaopheniidae), and cross-amplification in closely related species. *Biochem. Syst. Ecol.* 62, 137–141.
- Postaire, B., Magalon, H., Bourmaud, C.A.F., Gravier-Bonnet, N., Bruggemann, J.H., 2015c. Phylogenetic relationships within Aglaopheniidae (Cnidaria, Hydrozoa) reveal unexpected generic diversity. *Zool. Scripta.* 45 (1), 103–114.
- Powell, J.R., 2012. Accounting for uncertainty in species delineation during the analysis of environmental DNA sequence data. *Methods Ecol. Evol.* 3, 1–11.
- Prada, C., McIlroy, S.E., Beltrán, D.M., Valint, D.J., Ford, S.A., Hellberg, M.E., Coffroth, M.A., 2014. Cryptic diversity hides host and habitat specialization in a gorgonian-algal symbiosis. *Mol. Ecol.* 23, 3330–3340.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Puillandre, N., Baylac, M., Boisselier-Dubayle, M.C., Cruaud, C., Samadi, S., 2009. An integrative approach to species delimitation in *Benthomangelia* (Mollusca: Conoidea). *Biol. J. Linn. Soc.* 96, 696–708.
- Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012a. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* 21, 1864–1877.
- Puillandre, N., Modica, M.V., Zhang, Y., Sirovich, L., Boisselier-Dubayle, M.C., Cruaud, C., Holford, M., Samadi, S., 2012b. Large-scale species delimitation method for hyperdiverse groups. *Mol. Ecol.* 21, 2671–2691.
- Puillandre, N., Macpherson, E., Lambourdière, J., Cruaud, C., Boisselier-Dubayle, M.C., Samadi, S., 2011. Barcoding type specimens helps to identify synonyms and an unnamed new species in *Eumunida* Smith, 1883 (Decapoda: Eumunidiidae). *Invert. Syst.* 25, 322.
- R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 <<http://www.R-project.org/>>.
- Rambaut, A., Drummond, A.J., 2013. TreeAnnotator v1.7.0. URL: <<http://beast.bio.ed.ac.uk/treeannotator>>.
- Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6. URL: <<http://beast.bio.ed.ac.uk/Tracer>>.
- Reid, N.M., Carstens, B.C., 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evol. Biol.* 12, 196.

- Reimer, J.D., Takishita, K., Ono, S., Maruyama, T., 2007. Diversity and evolution in the zoanthid genus *Palythoa* (Cnidaria: Hexacorallia) based on nuclear ITS-rDNA. *Coral Reefs* 26, 399–410.
- Ridgway, T., Sampayo, E.M., 2005. Population genetic status of the Western Indian Ocean: what do we know? *WIO J. Mar. Sci.* 4, 1–9.
- Ronowicz, M., Boissin, E., Schuchert, P., Postaire, B., Bourmaud, C.A.-F., Gravier-Bonnet, N., 2016. Modern alongside traditional taxonomy - integrative systematics of the genera *Gymnangium* Hincks, 1874 and *Taxella* Allman, 1874 (Hydrozoa, Aglaopheniidae). *PLoS ONE*, submitted for publication.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542.
- Rosenberg, N.A., 2004. Distruct: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4, 137–138.
- Samadi, S., Barberousse, A., 2006. The tree, the network, and the species. *Biol. J. Linn. Soc.* 89, 509–521.
- Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., Stauffer, C., Christian, E., Crozier, R.H., 2010. Integrative taxonomy: a multisource approach to exploring biodiversity. *Annu. Rev. Entomol.* 55, 421–438.
- Schuchert, P., 2014. High genetic diversity in the hydroid *Plumularia setacea*: a multitude of cryptic species or extensive population subdivision? *Mol. Phylogenet. Evol.* 76, 1–9.
- Schuchert, P., 2005. Species boundaries in the hydrozoan genus *Coryne*. *Mol. Phylogenet. Evol.* 36 (1), 194–199.
- Schuchert, P., 2003. Hydroids (Cnidaria, Hydrozoa) of the Danish expedition to the Kei Islands. *Steenstrupia* 27, 137–256.
- Shanks, A.L., 2009. Pelagic larval duration and dispersal distance revisited. *Biol. Bull.* 216, 373–385.
- Siegel, D.A., Mitarai, S., Costello, C.J., Gaines, S.D., Kendall, B.E., Warner, R.R., Winters, K.B., 2008. The stochastic nature of larval connectivity among nearshore marine populations. *Proc. Natl. Acad. Sci. USA* 105, 8974–8979.
- Souter, P., 2010. Hidden genetic diversity in a key model species of coral. *Mar. Biol.* 157, 875–885.
- Spalding, M., Fox, H.E., Allen, G.R., Davidson, N., Ferdeña, Z.A., Finlayson, M., Halpern, B.S., Jorge, M.A., Lombana, A.L., Lourie, S.A., Martin, K.D., McManus, E., Molnar, J., Recchia, C.A., Robertson, J., 2007. Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience* 57, 573–583.
- Stampar, S.N., Maronna, M.M., Vermeij, M.J.A., Silveira, F.L.D., Morandini, A.C., 2012. Evolutionary diversification of banded tube-dwelling anemones (Cnidaria; Ceriantharia; *Isarachnanthus*) in the Atlantic Ocean. *PLoS ONE* 7, e41091.
- Talavera, G., Dincă, V., Vila, R., 2013. Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods Ecol. Evol.* 4, 1101–1110.
- Teske, P.R., Papadopoulos, I., Zardi, G.I., McQuaid, C.D., Edkins, M.T., Griffiths, C.L., Barker, N.P., 2007. Implications of life history for genetic structure and migration rates of southern African coastal invertebrates: planktonic, abbreviated and direct development. *Mar. Biol.* 152, 697–711.
- Turini, F.G., Steinert, C., Heubl, G., Bringmann, G., 2014. Microsatellites facilitate species delimitation in Congolese *Ancistrocladus* (Ancistrocladaceae), a genus with pharmacologically potent naphthylisoquinoline alkaloids. *Taxon* 63 (2), 329–341.
- Vernooy, R., Haribabu, E., Muller, M.R., Vogel, J.H., 2010. Barcoding life to conserve biological diversity: beyond the taxonomic imperative. *Plos Biol.* 8 (7), e1000417.
- Weersing, K., Toonen, R.J., 2009. Population genetics, larval dispersal, and connectivity in marine systems. *Mar. Ecol. Prog. Ser.* 393, 1–12.
- Yule, G.U., 1925. A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, F.R.S. *Philos. Trans. R. Soc., B* 213, 21–87.
- Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29 (22), 2869–2876.