**Cryptic divergence and strong population structure in the colonial
invertebrate *Pycnoclavella communis* (Ascidacea) inferred from molecular
data**

Rocío Pérez-Portela & Xavier Turon

Department of Animal Biology (Invertebrates), Faculty of Biology, University of Barcelona,
645, Diagonal Avenue, 08028 Barcelona, Spain.

Corresponding author address:

Rocío Pérez-Portela

Department of Animal Biology (Invertebrates), Faculty of Biology, University of Barcelona,
645, Diagonal Avenue, 08028 Barcelona, Spain

Phone number: +(34) 93 4021441

Fax number: +(34) 93 4035740

e-mail: rocio_perez@ub.edu

Number of text pages: 24

Number of Figures: 3

Number of Tables: 9

Number of Appendixes: 2

Abstract

We studied sequence variation in the mitochondrial gene COI for 135 individuals from eight Mediterranean populations of the colonial ascidian *Pycnoclavella communis* across most of its presently known range of distribution in the Mediterranean. Three haplotypes from Atlantic locations were also included in the study. Phylogenetic, phylogeographic, and population genetic analyses were used to unravel the genetic variability within and between populations. The study revealed 32 haplotypes for COI, 29 of them grouped within two Mediterranean lineages of *P. communis* (mean nucleotide divergence between lineages was 8.55%). Phylogenetic and network analyses suggest the possible existence of cryptic species corresponding to these two lineages. Population genetic analyses were restricted to the five populations belonging to the main genetic lineage, and for these localities we compared the information gleaned from COI sequence data and from eight microsatellite loci. A high genetic divergence between populations was substantiated using both kinds of markers (COI, global $F_{st} = 0.343$, microsatellite loci, global $F_{st} = 0.362$). There were high numbers of private haplotypes (COI) and alleles (microsatellites) in the populations studied. Restricted gene flow and inbreeding occur in the present range of distribution of the species. Microsatellite loci showed a strong incidence of failed amplifications, which we attribute to the marked intraspecies variability that hampered the application of these highly specific markers. Our results show important genetic variability at all levels studied, from within populations to between basins, possibly coupled to speciation processes. This variability is attributable to restricted gene flow among populations due to short distance dispersal of the larvae.

Keywords: free swimming larvae, restricted dispersal, COI, microsatellites, western Mediterranean

Introduction

Understanding the patterns of genetic variability of benthic invertebrates is crucial to assess speciation processes, connectivity between populations, susceptibility to local disturbance, and management guidelines (Shanks et al. 2003; Palumbi, 1994, 2004; Kinlan & Gaines 2003). Populations of marine species may have strong genetic structure even in groups with a supposedly high dispersal potential (Palumbi et al., 1997, 2004; Hellberg, 1994, 1996), resulting from discontinuities in gene flow as a consequence of physical and biological processes. Indeed, much cryptic genetic variation in marine habitats has been uncovered by the advent of molecular techniques (Knowlton, 2000; Féral, 2002). Different reproductive strategies in modular organisms have the potential to influence the genetic structure of the populations (Jackson, 1986; Hellberg 1994, 1996; Ayre et al., 1997) and such patterns of genetic variation among populations can be used to infer gene flow and the evolutionary history of the species.

Advances in the development of molecular markers using DNA have provided a wide range of tools for estimating patterns of relatedness and connectivity among benthic populations, understanding the spatial distribution of alleles and analyzing the populations' structure and demographic history (Emerson et al., 2001; Waples & Gaggiotti, 2006). Mitochondrial DNA sequence data have been used in a large number of studies for inferring the evolutionary and demographic past of both populations and species (Avise et al., 1987; Avise, 2000; Palumbi et al., 1997). Its genetic variability is most appropriate for phylogeny, phylogeography and population genetic approaches in which analyses of mtDNA are used to date events in the history of species, on the assumption that divergence within species is correlated with time (Moritz et al., 1987). However, a matrilineal phylogeny inferred from mitochondrial data gives us only a fraction of the genealogical information (Avise, 1998) and, for this reason, it is advisable to combine the information from mitochondrial markers with nuclear ones in order to obtain a more comprehensive picture of phylogeographical patterns (Ballard & Whitlock, 2004).

Microsatellite markers, on the other hand, have been widely employed in ecological and evolutionary studies over the past two decades. Because of their neutrality, codominance and high polymorphism, they are a powerful tool for inferring processes at the population level (Jarne & Lagoda, 1996; Zane et al., 2002). For example, analyses with the sponge *Crambe crambe* revealed homogeneity for the COI gene between distant populations but high

levels of polymorphism in microsatellite loci (Duran et al., 2002; Duran et al., 2004a). However, the marked specificity of microsatellites and the frequent occurrence of null alleles often hamper the application of these molecular tools (Dakin & Avise, 2004). Some studies have shown that amplification success is, in general, inversely correlated with molecular distances (Estoup et al., 1995; Jarne & Lagoda, 1996). The correlation between molecular divergence and microsatellite amplification may limit the usefulness of these markers when intraspecies genetic divergence is high.

Although modular marine invertebrates, with complex life cycles and asexual reproduction, are good candidates for the study of genetic variation at several scales using microsatellites (Calderon et al., 2007), most studies to date have been accomplished using allozymes (e.g., McFadden, 1997; Yund & O'Neil, 2000; Lazoski et al., 2001; Whalan et al., 2005). In particular, there are only three species of colonial ascidians in which microsatellite markers have been isolated (Pancer et al., 1994; Stoner et al., 1997; Maclean et al., 2004; Pérez-Portela et al., 2006).

Colonial ascidians are modular invertebrates widely distributed in marine systems, whose complex life cycles include sexual and asexual reproduction events, colonial fusion and chimerism (Sommerfeldt et al., 2003; Rinkevich, 2005). Most colonial species are brooders and the larvae have a restricted dispersal capability, since planktonic life span varies from a few minutes to hours (Millar, 1971; Svane & Young, 1989; Davis & Butler 1989), restricting connectivity between geographically close populations (Ayre et al., 1997). Frequent asexual reproduction and short dispersal distances led to the prediction that populations should be maintained by highly localised dispersal and a high degree of relatedness should exist at local scales (Jackson, 1986; Grosberg 1987; Ayre et al., 1997; Yund & O'Neil, 2000). For instance, populations of the compound species *Botryllus schlosseri* are characterised by deficiencies of heterozygotes, a feature characteristic of species with a considerable level of inbreeding (Stoner et al., 1997; Ben-Shlomo et al., 2001; Stoner et al., 2002). However, post-larval dispersal may increase substantially the dispersal capabilities of the species. For instance, rafting can have an important role in long range dispersal of colonial ascidian species (Worcester 1984; Muniz et al., 2006). Human-mediated transport is also a common mechanism enhancing dispersal of marine invertebrates (Carlton 2003). These mechanisms, however, are especially likely to affect species settling on drift material such as sea grass blades or seaweeds, or in artificial substrata (ship hulls, platforms).

Our study is focused on the recently described colonial ascidian *Pycnoclavella communis* Pérez-Portela et al., 2007, which is common but patchily distributed on open rocky

littoral habitats along the eastern Atlantic (Canary Islands and Madeira) and western Mediterranean (Pérez-Portela et al., 2007). This is one of the rare ascidian species for which microsatellite markers have been identified (Pérez-Portela et al., 2006). Although several colour morphs were described in the Mediterranean, a phylogenetic study of the genus (Pérez-Portela et al., 2007) showed that this species featured a high degree of genetic variation which was not correlated with differences in colour. The goal of the present study is to analyse the genetic structure of this variable species in the western Mediterranean, using phylogenetic inference, network estimation and population genetics, in order to explore geographical patterns and evolutionary events that can explain the distribution of genetic diversity of this species. To this end, we applied two different sets of molecular markers, mtDNA and microsatellites, to explore the degree of resolution of these different markers at the intraspecific level. As the species lives attached on rocky substrates and is not present in marinas or other artificial structures, we do not expect that post-larval processes have a noticeable influence on its dispersal ranges. We hypothesize, therefore, that genetic variation will be found at different scales, and that genetic differentiation can build up even between geographically close populations.

Materials and Methods

Sampling and DNA extraction

Individuals morphologically attributable to the yellow morph of *Pycnoclavella communis* were collected from eight western Mediterranean localities within two areas of the Spanish littoral covering most of the presently known geographical range of the species in the Mediterranean: Palamós 41° 5' N 2° 08' E (P), Blanes 41° 40' N 2° 48.2' E (BL) and Tossa 41° 43.2' N 2° 56.4' E (T) (northern Spanish littoral) and Punta de la Mona 36° 43.4' N 3° 08' W (PM), Cerro Gordo 36° 43.15' N 2° 08' W (CG), Cabo de Palos 37° 37.9' N 0° 41.5' W (CP), Carboneras 36° 59.7' N 1° 53.3' W (CA), and Aguamarga 36° 54' N 1° 58.8' W (AG) (southern Spanish littoral) (Fig. 1). With these sampling localities we had populations separated by a wide range of distances. We considered *a priori* each of these localities as distinct populations based on the assumed short dispersal range of the larvae. Colonial ascidians are among the rare groups for which direct measurements on larval dispersal have been made, and the results have usually found greatly restricted dispersal, of the order of a few metres (Olson 1987;

Worcester, 1994). In a related clavelinid (*Podoclavella*), Davis & Butler (1989) found that almost 80% of the larvae settled within 2.5m of the parent colony. This extremely short range dispersal lends support to an expectation of low number of migrants over generations even among populations separated by a few tens of kilometres.

All the colonies included in this study (see below) were sampled by scuba diving between 5 m and 30 m deep and separated at least 2 m from each other in order to minimize the chances of sampling the same clone (in this stolonial species, fission phenomena can lead to the formation of clones at small distance ranges). The specimens were identified following Pérez-Portela et al. (2007) and immediately preserved in absolute ethanol and kept at -20°C until processed. No morphological differences could be observed between the specimens from the different populations studied. Once in the laboratory, a few zooids of each preserved colony were separated with forceps for DNA extraction. Total DNA was extracted for mitochondrial DNA sequencing and microsatellite amplification using the protocol described in Pascual et al. (1997).

(-----Figure 1 approximately here -----)

Mitochondrial DNA sequencing

Sequences of a fragment of the cytochrome *c* oxidase subunit I (COI) mitochondrial gene were obtained from 135 specimens from the eight populations studied. Universal primers HCO2198 & LCO1490 (Folmer et al., 1994) and specific primers PING & PONG (Pérez-Portela et al., 2007) were used. Amplification was performed in a 20µl total reaction volume with 0.5 µl of each primer (25 µM), 0.5 µl dNTP's (10 mM), 2µl 10X buffer, 1.6 µl MgCl₂ (Promega), 0.2µl Taq polymerase (Promega) and 0.5 µl template DNA. A single soak at 94°C for 2 min was followed by 35 cycles (denaturation at 94°C for 1 min, annealing at 40-50 °C for 1 min, and extension at 72°C for 1.5 min) and a final extension at 72°C for 7 min, on a Perkin Elmer 9600 DNA amplifier. The same primers were used for the sequencing reaction, and the PCR products were sequenced with an ABI Big-Dye Ready-Reaction Perkin Elmer kit on an ABI Prism 377XL automated sequencer (Scientific and Technical Services of the University of Barcelona). Sequences were edited and aligned using the Bioedit Sequence Alignment Editor (Hall, 1999) and alignment was confirmed by eye. No gap was needed in

the alignment and all sequences could be translated into amino acids without stop codons. The nucleotide sequences obtained in this study have been deposited in EMBL (pending).

Microsatellite amplification and genotyping

Microsatellite genotyping was performed on some populations of the northern Mediterranean area (Blanes and Tossa) and of the southern zone (Punta de la Mona and Cerro Gordo). A total of 110 individuals from these sampling sites was screened for variation at each of eight polymorphic microsatellite loci previously isolated and described for *Pycnoclavella communis* (the *Pycnoclavella* sp. in Pérez-Portela et al., 2006 is *P. communis*). Forward primers were 5' labelled with a fluorescent dye as in Pérez-Portela et al. (2006) and PCR was performed under the conditions described in that work. Allele sizes were estimated on an automated sequencer ABI Prism-3700 relative to an internal standard (EcoGen 70-400). Alleles were visualised and determined with GENESCAN and GENOTYPER software.

Phylogenetic analyses of COI gene

The phylogenetic information in the dataset of the 135 Mediterranean specimens was studied. Three Atlantic haplotypes of *Pycnoclavella communis* obtained from Pérez-Portela et al. (2007) were also included in the analyses to assess the relationship between Atlantic and Mediterranean haplotypes (EMBL AM403665, AM403666 and AM403667). Moreover, one sequence of *Pycnoclavella aurilucens* (EMBL AM403692) was used as outgroup. The best-fit model of nucleotide substitution for our data was selected by statistical comparison of 56 different models of evolution with the Modeltest 3.0 program (Posada & Crandall, 1998), using the Akaike Information Criterion (AIC). The model selected was then input into the Treefinder program (Jobb et al., 2004) and a haplotype tree was estimated. Nodal support was assessed with 1,000 bootstrap replicates.

Haplotype network of COI gene

The relationship between haplotypes, under the null hypothesis of no genetic differentiation among populations of the whole area studied, was assessed by an unrooted network. We used the Network program (<http://www.fluxus-engineering.com/sharenet.htm>), which assumes the

median-joining network method in the absence of recombination (Bandelt et al., 1999). This method begins by combining the Minimum-Spanning Trees within a single network. With a parsimony criterion, median vectors (which represent missing intermediates) are added to the network.

Population genetics and demographic analysis of COI and microsatellite data

Due to a high divergence between two lineages of *Pycnoclavella communis* found within the Mediterranean basin and to avoid an incorrect lumping of separate gene pools (perhaps separate species, see Discussion), we applied the population genetics analyses only to the five populations belonging to the largest genetic clade found (i.e., Palamós, Tossa, Blanes, Punta de la Mona and Cerro Gordo) which included the zone where *P. communis* was described.

COI gene

Nucleotide diversity (π) and haplotype diversity (h) (Nei, 1987) were estimated using ARLEQUIN version 3.01 (Schneider et al., 2000). The same program was used to calculate the pairwise genetic distances (F_{st}) among populations. Their significance was calculated by performing 10,000 permutations of the dataset. We also performed analyses of molecular variance (AMOVA) to examine hierarchical population structure by pooling the populations in southern and northern Mediterranean groups, as well as without grouping populations. We executed 16,000 permutations to guarantee having less than 1 % difference with the exact probability in 99% of cases and used our prior expectation of a genetic division between northern and southern populations. An exact test of population differentiation based on haplotype frequencies (Raymond & Rousset, 1995) was performed to test the null hypothesis that observed haplotype distribution is random with respect to sampling location. The significance of individual tests was estimated by comparison with a simulated distribution constructed from 10,000 permutations of the original matrix. Isolation by distance was assessed from the correlation between matrices of pairwise $F_{st}/(1-F_{st})$ values and the logarithm of the geographical distances between populations (Mantel test, 10,000 permutations, Rousset, 1997). Geographical distance between populations was estimated as minimum distances in kilometres along the coastal line.

To determine the history of effective population size we computed Tajima's D -test (Tajima, 1989), Fu's F_s test (Fu, 1997) and the R_2 test (Ramos-Onsins & Rozas, 2002) using the DnaSP 4.10 program (Rozas et al., 2003). All analyses were performed both on the

separate populations and on the populations pooled into northern and southern geographic groups. Fu's F_s and R_2 have been shown to perform best at detecting population growth (Ramos-Onsins & Rozas, 2002).

Microsatellite loci

We used microsatellite data for four of the five populations considered. The fifth population (Palamós) could not be included in the analyses since most of the individuals and loci failed in the amplification. All computations were performed using the software packages GENEPOP 3.3 (Raymond & Rousset, 1995) and ARLEQUIN version 3.01 (Schneider et al., 2000). Allele frequencies, number of observed and expected heterozygotes and F_{is} (Weir & Cockerham, 1984) were estimated for each locus and location. The exact tests for deviation from Hardy-Weinberg expectation at each locus and for genotypic linkage disequilibria among all pairs of loci and for all locations were calculated. The genetic differentiation among the sampled sites was quantified assuming an infinite alleles model (IAM) (Kimura & Crow, 1964) using F_{st} (Weir & Cockerham, 1984), and the significance of pairwise values was calculated by performing 10,000 permutations of the data. The IAM assumes that each mutation results in a new, previously undefined allele. We also quantified the genetic differentiation under a stepwise mutation model (SMM) for microsatellite loci with Rho_{st} (Michalakis & Excoffier, 1996). The SMM model assumes that all microsatellite mutations occur as an addition or loss of a single repeat unit from a pre-existing allele. Genetic differentiation is, therefore, estimated as a function of the sum of squared differences in repeat numbers between alleles. An AMOVA was also performed both pooling the populations in northern and southern Mediterranean groups, and without grouping populations. A Mantel test procedure (10,000 permutations) was performed to assess isolation by distance.

Results

Phylogenetic analyses of COI

The final length of the COI sequences after alignment and trimming was 546 bp. A total of 32 different haplotypes were obtained from 138 specimens, collected along the western Mediterranean and eastern Atlantic coasts (see Appendix I), indicating a high degree of polymorphism. Overall, there were 92 polymorphic sites (16.8 %), and nine substitutions

resulted in non-synonymous changes. Nucleotide variation was scattered across the entire sequenced region and was mainly restricted to the third base position of the codons (82%). In the Modeltest procedure, the best model among those evaluated by the Akaike Information Criterion (AIC) was TVM+I+G. The parameters of the model were as follows: base frequencies, A = 0.2599, C = 0.1359, G = 0.1841, T = 0.4200; substitution rate matrix, T-C = 255.3235, T-A = 19.8910, T-G = 1, C-A = 32.1607, C-G = 10.9353, A-G = 255.3237; proportion of invariable sites (I) = 0.5015; and gamma distribution shape parameter (G) = 0.5650. These values were input into Treefinder, except for the proportion of invariable sites, not available in this program. In addition, we used different change rates according to codon position. The rates estimated by the program were: first position = 0.4638, second position = 0.0729, and third position = 2.4630. A maximum likelihood tree (ML) tree was obtained under this model (Fig. 2).

Two main mtDNA lineages, named lineage A and B for convenience, were inferred and appeared well supported (bootstrap values > 90 %). These major mtDNA lineages did not correspond to two distinct geographic regions. Lineage A grouped 22 haplotypes from five Mediterranean populations of the species *Pycnoclavella communis* located in the northern and southern ends of the Mediterranean range studied, together with a clade formed by the haplotypes from the Atlantic samples. Lineage B, on the other hand, included seven haplotypes from three geographically intermediate (between the North and South Mediterranean) populations (Carboneras, Aguamarga and Cabo de Palos). The main parameters describing variability within the clades and populations, namely the number of polymorphic sites, number of haplotypes, haplotype diversity (h) and nucleotide diversity (π), are summarised in Table 1. The number of haplotypes per site ranged from three to six and there was a trend towards a higher number of haplotypes in northern Mediterranean populations. It was found that 85% of haplotypes were private for one or another population. Mean nucleotide divergence between the two lineages was high (8.55 %), and there were no shared haplotypes among them. Genetic differentiation was also found within lineages. In particular, there was a 6% nucleotide divergence between Mediterranean and Atlantic populations within the Lineage A. One non-synonymous change separated Lineages A and B, and another separated Atlantic and Mediterranean samples of Lineage A. However, most replacement changes occurred within populations (Fig. 3).

(-----Figure 2 approximately here -----)

(-----Table 1 approximately here -----)

Phylogeography and demography

The haplotype network cladogram is presented in Figure 3. Most of the haplotypes analysed fell into groups corresponding with the results obtained in the ML tree. The Mediterranean clade of Lineage A was separated from Lineage B by more than 35 mutational steps in the network. On the other hand, in Lineage A the Mediterranean clade was differentiated from the Atlantic haplotypes by more than 20 mutational steps, and there was also a marked separation within the Atlantic haplotypes between the Canary Islands and Madeira. Within the groups found, the number of mutations ranged from one to four between haplotypes (with one exception: haplotype H10 of the Mediterranean clade of Lineage A).

(-----Figure 3 approximately here -----)

The results of the neutrality tests are shown in Table 2. We analysed the two Mediterranean clades and their separate populations. Ramos-Onsins & Rozas' R_2 test and Fu's F_s test were not significant in any case, while Tajima's D test was significant for the population of Punta de la Mona, and also produced a significant outcome for the whole Mediterranean clade of Lineage A. Overall, therefore, the majority of tests were not significant, indicating no evidence of selection acting on this locus and that no major expansion has occurred recently in the areas studied.

(-----Table 2 approximately here -----)

Population genetics

Given the results of the phylogenetic analysis and the haplotype network analysis, we performed further population genetics analyses only in the Mediterranean clade of Lineage A to avoid the confounding effect of mixing excessively distant lineages.

COI sequences

The five populations analysed had 22 haplotypes and most of them (90.9 %) were private. Only two haplotypes were shared between two populations each (see Appendix I). We found high haplotype diversity (0.858 ± 0.00063 , mean $\pm$ SE) and low values of nucleotide diversity (0.00982 ± 0.00099) per population (Table 1). The pairwise tests for genetic differentiation among these Mediterranean populations of *Pycnoclavella communis* based on F_{st} values revealed moderate to strong differentiation between locations (global F_{st} of 0.343). All comparisons, except those between Punta de la Mona and Cerro Gordo, displayed significantly higher genetic divergence than would be expected in a panmictic population (Table 3). The results of the exact test for population differentiation based on haplotype frequencies between locations were similar but, in this case, all comparisons were significant (data not shown).

(-----Table 3 approximately here -----)

The AMOVA analysis pooling the five sampling sites according to geographical origin (northern vs southern Spanish Mediterranean coast) revealed that the genetic variance component associated with geographical origin was high (26.61 %) but non-significant. Most of the variance observed (58.88 %) was concentrated within populations and 14.51% of variance was related to between-population differentiation. The latter two components were significant (Table 4). The AMOVA performed without grouping sampling sites again revealed a strong variance component associated with within-population differentiation (65.64 %) and a significant genetic variance related to between-population differentiation (34.36 %). AMOVA and pairwise tests revealed, therefore, significant genetic differentiation both within and between populations in the area studied. Although the variance associated with the North-South geographical distinction was not significant, the Mantel test showed a significant correlation between $F_{st}/(1-F_{st})$ and geographical distance (Mantel test, $P= 0.038$), reflecting some degree of isolation by distance.

(-----Table 4 approximately here -----)

Microsatellite loci

55 alleles were found in the eight microsatellite loci analysed in four populations. The number of alleles per locus ranged from two to 11 (see Appendix II). Allele frequency distribution

across the eight polymorphic loci did not show the same pattern in different populations. There were private alleles in all populations studied and many alleles were not shared even between adjacent populations (e.g. all alleles of PYC-35 were private). No significant linkage disequilibrium between pairs of loci was detected across the populations. All populations analysed showed deviation from Hardy-Weinberg equilibrium ($P < 0.01$) and a heterozygote deficiency (Table 5). Failure of amplification was frequent in all loci and populations (Table 6). In all cases, we tried DNA re-extraction, amplification under less stringent reaction conditions, and re-design of primers, but these efforts did not result in any significant improvement of genotype readings.

(-----Table 5 approximately here -----)

(-----Table 6 approximately here -----)

The comparison of multilocus F_{st} and Rho_{st} indicated a high level of differentiation between populations (Table 7). The pairwise values were significant for all pairs of populations. The different loci did not contribute equally to the inter-population differentiation, with single locus F_{st} and Rho_{st} ranging from 0.145 to 0.563 and from 0.018 to 0.747, respectively (Table 8).

(-----Table 7 and 8 approximately here -----)

The AMOVA analysis pooling the populations again showed no significant contribution of the geographical area (North vs South) to the pattern of genetic variance, while ca 38% of variation was found among populations and ca 64% of the genetic variance was detected within populations, both components being highly significant (Table 9). The AMOVA performed without groupings again revealed a significant value of the genetic variance associated with among- and within-population differentiation. There was no evidence of isolation by distance between geographical locations using the microsatellite dataset (Mantel test, $P < 0.05$).

(-----Table 9 approximately here -----)

Discussion

The phylogenetic tree and the haplotype network revealed significant genetic differences between the studied populations of *Pycnoclavella communis*, with the two main lineages found being separated by 8.55 % sequence divergence in the COI gene. Molecular data have shown a prevalence of cryptic and sibling species in marine environments that had gone undetected in previous morphological studies (Palumbi, 1994; Knowlton, 2000; Féral, 2002), but the criteria used to distinguish species based on molecular data are controversial (Avise, 1994; Templeton, 2001; Sites & Marshall, 2003; Lee, 2004). In particular, the degree of divergence among species in the COI gene is diverse among and within taxonomic groups, and there is no single threshold value (Hebert et al., 2003). We have several data on COI variation in colonial ascidians for comparison: cryptic species of *Clavelina lepadiformis* had genetic differences of ca. 5 % nucleotide divergence for COI, but ecophenotypic differences were recorded as well (Tarjuelo et al., 2001; DeCaralt et al., 2002; Turon et al., 2003). For the cosmopolitan ascidian *Botryllus schlosseri*, the mean number of nucleotide differences between COI haplotypes was 5.45 % and there was no evidence that there were different species (López-Legentil et al., 2006). In another colonial ascidian, *Pseudodistoma crucigaster*, colour morphs featuring 2.12 % COI sequence divergence were suggested to be separate species (Tarjuelo et al., 2004). In the genus *Cystodytes*, two groups of haplotypes separated by at least 9.9 % sequence divergence were taken as belonging to different species (López-Legentil & Turon, 2006). Thus, the threshold separating intraspecies from interspecies divergence for this gene seems to be variable within the group of colonial ascidians, and complementary biological studies were needed to support genetic results (e.g., DeCaralt et al., 2002; López-Legentil et al., 2005). Detailed morphological and biological studies should, therefore, complement our genetic data to ascertain whether we are dealing with a highly variable species or with a group of cryptic species.

Several studies have reported genetic discontinuity between Atlantic Ocean and Mediterranean Sea in species with Atlanto-Mediterranean distribution (e.g. Borsa et al., 1997; Turon et al., 2003; Roman & Palumbi, 2004; Baus et al., 2005). The existence of a surface current of North Atlantic water reaching as far as Alboran Sea might mark the main barrier between Atlantic and Mediterranean populations for some marine invertebrates (Maldonado & Uriz, 1995). In our case, however, the populations of *Pycnoclavella communis* split into

two distinct mitochondrial clades within the Mediterranean basin, where divergence cannot be explained by macrogeographic features. One of these clades belongs to the same group as the Atlantic samples included for comparison (Lineage A), while the other clade (Lineage B) is exclusively Mediterranean and is interspersed between Mediterranean populations of Lineage A, which were found at both sides of the Alboran front. *Pycnoclavella communis* was formally described from samples included in Lineage A (Pérez-Portela et al., 2007).

In order to avoid the confounding effect of mixing two separate genetic clades (perhaps different species), we performed the population-level analyses only on the Mediterranean populations included in Lineage A. Both mitochondrial sequences and microsatellite data revealed a strong degree of genetic structure between the populations studied. This is reflected by significant F_{st} values in pairwise comparisons and a high percentage of private COI haplotypes and microsatellite alleles. Even geographically close populations, such as Blanes and Tossa, separated by only 10 km, were highly differentiated, with no shared haplotypes and 12 out of 19 microsatellite alleles unique to one or the other population. At the geographic scale considered, genetic divergence seems to increase with distance when using mitochondrial markers, but the outcome is not significant when using microsatellite data. Nonetheless, strong intrapopulation variance also exists, as revealed by AMOVA analyses using both sets of markers. The pattern found is that of a high haplotype diversity coupled with low nucleotide diversity. This pattern is often related to expansions after a period of small effective population size (Avise et al., 1984), but in our case the neutrality tests do not show any clear signal of population expansion. In another colonial ascidian, *Botryllus schlosseri*, López-Legentil et al. (2006) found the opposite pattern of low haplotype diversity and high nucleotide diversity in South European waters, with a lack of intermediate haplotypes. They attributed this pattern to repeated introductions in the area of a few, well differentiated variants of this cosmopolitan species. In our case, most haplotypes within the groups delimited by phylogenetic and network analyses were separated by one or a few mutations. A relatively recent history of the species in the area and the lack of population connectivity may provide the explanation for the observed pattern of proliferation of private haplotypes separated by few nucleotide changes.

Several studies in colonial ascidians have provided evidence that localised dispersal and mechanisms for retention of larvae close to parental zooids determine their genetic structure, playing a major role in maintaining established populations (Svane & Young, 1989; Ayre et al., 1997; Tarjuelo et al., 2004). The realised dispersal of ascidian larvae has been shown to be much smaller than that which larval lifespans could potentially make possible

(Grosberg, 1987; Olson & McPherson, 1987; Davis & Butler, 1989), thus creating genetic structure even at scales of a few meters (Yund & O’Neil, 2000). The finding of private haplotypes and alleles in all populations and loci provides strong evidence that populations of *Pycnoclavella communis* are maintained through local recruitment, in contrast with groups whose larvae have higher dispersal capabilities and that display homogenous population structure over the same area of this study (e.g. sea urchins, Duran et al., 2004b). It seems, therefore, that each population of *P. communis* is independently evolving due to infrequent genetic exchange (Slatkin, 1987), which is insufficient to homogenize haplotypes and allele frequencies among populations.

Our data therefore suggest a scenario in which a species, as a result of highly restricted gene flow, accumulates mutations leading to fast separation of lineages. Speciation events may have already taken place, given the marked genetic separation of the Mediterranean clade that we called Lineage B. Why this clade should have become so divergent is not clear at present. One hypothesis may be that in other parts of the Mediterranean or the Atlantic the species has experienced strong divergence (it should be noted here that the Atlantic haplotypes of the Canary Island and Madeira are separated by at least 20 mutations) and that the lineage B is the result of a separate arrival to the area from a distant, already divergent population source, and is now experiencing its own genetic differentiation fuelled by a lack of connectivity. Both the level of nucleotide diversity and the haplotype diversity found were always lower for populations included in Lineage B, suggesting a recent arrival. The localities associated with Lineage B lie along a known hydrographic feature, the Almería-Oran front (Tintoré et al., 1988), and hence within a cyclonic gyre of the water masses (Hopkins, 1985), which may have reinforced its isolation and could explain why this clade appears interspersed within the range of distribution of the larger of the Mediterranean clade studied here.

The number of alleles found at the microsatellite loci ranged between 2 and 11 (mean of 6.63), which is a moderate value if we compare it with other colonial and solitary ascidians (Pancer et al., 1994; Stoner et al., 2002; Ben-Shlomo et al., 2001; Maclean et al., 2004; Dupont et al., 2006). Microsatellite loci in *Pycnoclavella communis* showed a deficit of heterozygotes within all populations. Heterozygote deficiency is common in marine populations and has been found in other brooding colonial ascidians (Grosberg, 1991; Ben-Shlomo et al., 2001; Yound & O’Neil, 2000). It can be a result of several non-exclusive factors: inbreeding, self-fertilization, Wahlund effect or the presence of non-amplifying alleles (null alleles) (Whalan et al., 2005). A conspicuous feature of our dataset is the prevalence of non-amplifying alleles. This fact can explain the heterozygote deficiency found.

533 However, we cannot rule out completely other factors, as the heterozygote deficiency was
534 significant even when most of the individuals amplified.

535 Low dispersal capability can create patches of related individuals (substructuring) and
536 favour mating among relatives, and hence inbreeding. Self-fertilization cannot be rejected
537 either. *P. communis* is protandric but a monitoring of the reproductive cycle showed that there
538 is an overlap between male and female periods (Pérez-Portela, 2006) that might allow cross-
539 fertilization between individuals of the same colony. There was no evidence, such as repeated
540 multilocus genotypes or the preponderance of a few haplotypes, that asexual reproduction
541 played an important role at the population level. This outcome is most likely a result of our
542 sampling scheme designed to avoid sampling clones (individuals were separated by more than
543 2 m), rather than a true lack of clonal reproduction in this species.

544 Results obtained from the microsatellite loci in the present study should be taken with
545 caution due to null alleles, although it is reassuring in this sense that they are coincident with
546 conclusions drawn from mitochondrial sequence data. It is unclear why our microsatellite
547 dataset has been so plagued with null alleles (45 % of failed amplifications for the whole
548 dataset). Given our repeated attempts using different amplification conditions and primers
549 (when we had enough flanking sequence to redesign primers), we believe that our failures to
550 amplify indicate true null alleles and not technical problems. Null alleles originate when
551 mutations in the flanking region prevent the primers from annealing. Our species may be too
552 variable for these highly specific markers. This finding highlights a shortcoming of the
553 microsatellite technique when applied to species with high intraspecific variability (the mean
554 of nucleotide divergence between populations within the Mediterranean clade of Lineage A
555 was close to 2%).

556 In conclusion, our results show that *Pycnoclavella communis* features genetic
557 differentiation at several levels (between lineages, between Atlantic and Mediterranean basins
558 in Lineage A, between populations and within populations). If genetic differences between the
559 main lineages are a result of local ongoing processes or of introductions of already
560 differentiated clades could not be resolved. Future studies including more Atlantic and
561 Mediterranean populations could clarify their geographical origins. Whether Lineage B is
562 already reproductively isolated from the other Mediterranean clade, or both of them from the
563 Atlantic forms, also remains unclear at present, but the evidence for highly restricted gene
564 flow even in geographically close populations indicates that genetic diversity is building up.
565 The specific richness of *Pycnoclavella* recently uncovered in Southern Europe (Pérez-Portela
566 et al., 2007) may fall short of describing the true patterns of genetic variation of the genus, in

yet another instance of the difficulty of reconciling the static view of species definition with the dynamic processes that create them.

Acknowledgments

We thank Dr Cruz Palacín for her help with underwater work and Dr. Marta Pascual for her help with the microsatellite technical problems. The research was funded by projects REN2001-2312-C03-02 and CTM2004-05265-C02-01 of the Spanish Government.

References

- Avice, J.C., 1994. Molecular markers, natural history and evolution. Chapman & Hall, New York.
- Avice, J.C., 1998. The history and purview of phylogeography: a personal reflection. *Mol. Ecol.* 7, 371-379.
- Avice, J.C., 2000. Phylogeography. The history and formation of species. Harvard University Press, Cambridge, Massachusetts.
- Avice, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C., 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18, 489-522.
- Avice, J.C., Neigel, J.E., Arnold, J., 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Mol. Evol.* 20, 99-105.
- Ayre, D.J., Davis, A.R., Billingham, M., Llorens, T., Styan, C., 1997. Genetic evidence for contrasting patterns of dispersal in solitary and colonial ascidians. *Mar. Biol.* 130, 51-61.
- Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. *Mol. Ecol.* 13, 729-744.
- Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37-48.

597 Baus, E., Darrock, J., Bruford, M.W., 2005. Gene flow patterns in Atlantic and Mediterranean
598 populations of the Lusitanian sea star *Asterina gibbosa*. Mol. Ecol. 14, 3373-3382.

599 Ben-Shlomo, R., Douck, J., Rinkevich, B., 2001. Heterozygote deficiency and chimerism in
600 remote populations of a colonial ascidian from New Zealand. Mar. Ecol. Prog. Ser. 109, 109-
601 117.

602 Borsa, P., Naciri, M., Bahri, L., Chikhi, L., Garcia de León, F.J., Kotoulas, G., Bonhomme,
603 F., 1997. Intraspecific zoogeography of the Mediterranean: population genetic analysis on
604 sixteen Atlanto-Mediterranean species (fishes and invertebrates). Vie Milieu 47, 295-305.

605 Calderon, I., Ortega, N., Duran, S., Becerro, M.A., Pascual, M., Turon, X. (2007). Finding the
606 relevant scale: clonality and small-scale genetic structure in *Crambe crambe* (Porifera). Mol.
607 Ecol. 16: 1799-1810.

608 Carlton, J.T., 2003. Community assembly and historical biogeography in the North Atlantic
609 Ocean: the potential role of human-mediated dispersal vectors. Hydrobiologia 503, 1-8.

610 Dakin, E.E., Avise, J.C., 2004. Microsatellite null alleles in parentage analysis. Heredity 93,
611 504-509.

612 Davis, A.R., Butler, A.J., 1989. Direct observations of larval dispersal in the colonial ascidian
613 *Podoclavella moluccensis* Sluiter: evidence for closed populations. J. Exp. Mar. Biol. Ecol.
614 127, 189-203.

615 De Caralt, S., López-Legentil, S., Tarjuelo, I., Uriz, M.J., Turon, X., 2002. Contrasting
616 biological traits of *Clavelina lepadiformis* (Ascidacea) populations from inside and outside
617 harbours in the western Mediterranean. Mar. Ecol. Prog. Ser. 244, 125-137.

618 Dupont, L., Viard, F., Bishop, J.D.D., 2006. Isolation and characterization of twelve
619 polymorphic microsatellite markers for the invasive ascidian *Styela clava* (Tunicata). Mol.
620 Ecol. Notes 6, 101-103.

621 Duran, S., Pascual, M., Estoup, A., Turon, X., 2002. Polymorphic microsatellite loci in the
622 sponge *Crambe crambe* (Porifera: Poecilosclerida) and their variation in two distant
623 populations. Mol. Ecol. Notes 2, 478-480.

624 Duran, S., Pascual, M., Estoup, A., Turon, X., 2004a. Strong population structure in the
625 marine sponge *Crambe crambe* (Poecilosclerida) as revealed by microsatellite markers. Mol.
626 Ecol. 13, 511-522.

627 Duran, S., Palacín, C., Becerro, M.A., Turon, X., 2004b. Genetic diversity and population
628 structure of the commercially harvested sea urchin *Paracentrotus lividus* (Echinodermata,
629 Echinoidea). Mol. Ecol. 13, 3317-3328.

630 Emerson, B.C., Paradis, E., Thébaud, C., 2001. Revealing the demographic histories of
631 species using DNA sequences. Trends Ecol. Evol. 16, 707-716.

632 Estoup, A., Jarne, P., Cornuet, J.M., Solignac, M., 1995. Size homoplasy and mutational
633 process of interrupted microsatellite in two bee species *Apis mellifera* and *Bombus terrestris*
634 (*Apiae*). Mol. Biol. Evol. 12, 1074-1084.

635 Féral, J.P., 2002. How useful are the genetic markers in attempts to understand and manage
636 marine biodiversity?. J Exp. Mar. Biol. Ecol. 268, 121-145.

637 Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for
638 amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan
639 invertebrates. Mol. Mar. Biol. Biotech. 3, 294-299.

640 Fu, Y.X., 1997. Statistical tests of neutrality against population growth, hitchhiking and
641 background selection. Genetics 147, 915–925.

642

643 Grosberg, R.K., 1987. Limited dispersal and proximity-dependent mating success in the
644 colonial ascidian *Botryllus schlosseri*. Evolution 41, 372-384.

645 Grosberg, R.K., 1991. Sperm-mediated gene flow and the genetic structure of a population of
646 the colonial ascidians *Botryllus schlosseri*. Evolution 45, 130-142.

647 Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis
648 program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.

649 Hebert et al. 2003

650 Hellberg, M.E., 1994. Relationships between inferred levels of gene flow and geographic
651 distance in a philopatric coral, *Balanophyllia elegans*. Evolution 48, 1829-1854.

652 Hellberg, M.E., 1996. Dependence of gene flow on geographic distance in two solitary corals
653 with different larval dispersion capabilities. *Evolution* 50, 1167-1175.

654 Hopkins, T.S., 1985. Physics of the sea. In: Western Mediterranean. Pergamon Press, Oxford.

655 Jackson, J.B.C., 1986. Modes of dispersal of clonal benthic invertebrates: Consequences for
656 species distributions and genetic structure of local populations. *B. Mar. Sci.* 39, 588-606.

657 Jarne, P., Lagoda, P.J.L., 1996. Microsatellites, from molecules to populations and back.
658 *Trends Ecol. Evol.* 11, 424-429.

659 Jobb, G., Von Haeseler, A., Strimmer, K., 2004. TREEFINDER: a powerful graphical
660 analysis environment for molecular phylogenetics. *BMC Evolutionary Biology* 4, 18.

661 Kimura, M., Crow, J.F., 1964. The number of alleles that can be maintained in a finite
662 population. *Genetics* 49, 725-738.

663 Kinlan, B.P., Gaines, S.D., 2003. Propagule dispersal in marine and terrestrial environments:
664 a community perspective. *Ecology* 84: 2007-2020

665 Knowlton, N., 2000. Molecular genetic analyses of species boundaries in the sea.
666 *Hydrobiologia* 420:73-90

667 Lazoski, C., Solé-Cava, A.M., Boury-Esnault, N., Klautau, M., Russo, C.A.M., 2001. Cryptic
668 speciation in a high gene flow scenario in the oviparous marine sponge *Chondrosia*
669 *reniformis*. *Mar. Biol.* 139, 421-429.

670 Lee, M.S., 2004. The molecularisation of taxonomy. *Invertebr Syst* 18(1), 1-6

671 López-Legentil, S., Ruchty, M., Domènech, A., Turon, X., 2005. Life cycles and growth rates
672 of two morphotypes of *Cystodytes* (Ascidiacea) in the western Mediterranean. *Mar. Ecol.*
673 *Prog. Ser.* 296, 219-228.

674 López-Legentil, S., Turon, X., Planes, S., 2006. Genetic structure of the star sea squirt,
675 *Botryllus schlosseri*, introduced in southern European harbours. *Mol. Ecol.* 15, 3957-3967.

676 López-Legentil, S., Turon, X., 2006. Population genetics, phylogeography and speciation of
677 *Cystodytes* (Ascidiacea) in the Western Mediterranean Sea. *Biol. J. Linn. Soc.* 88, 203-214.

678 Maclean, L., Pardini, A.T., Noble, L.R., Bishop, J.D.D., Jones, C.S., 2004. Polymorphic
679 dinucleotide microsatellite loci in the clonal ascidian *Diplosoma listerianum*: predominance
680 of compound and highly interrupted imperfect repeats. Mol. Ecol. Notes 4, 283-285.

681 Maldonado, M., Uriz, M.J., 1995. Biotic affinities in a transitional zone between the Atlantic
682 and the Mediterranean: a biogeographical approach based on sponges. J Biogeogr, 22, 89-110.

683 McFadden, C.S., 1997. Contributions of sexual and asexual reproduction to population
684 structure in the clonal soft coral, *Alcyonium rudyi*. Evolution 51, 112-126.

685 Michalakis, Y., Excoffier, L., 1996. A genetic estimation of population subdivision using
686 distances between alleles with especial reference to microsatellite loci. Genetics 142, 1061-
687 1064.

688 Millar, R.H., 1971. The Biology of Ascidians. Adv. Mar. Biol. 9, 1-100.

689 Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of animal mitochondrial DNA:
690 Relevance for population biology and systematics. Ann. Rev Ecol Syst 18, 269-292.

691 Muniz, G., Lembo, L.F., Nisaka, V., 2006. Low genetic differentiation between isolated
692 populations of the colonial ascidian *Symplegma rubra* Monniot C., 1972. Mar. Biol. 148, 807-
693 815.

694 Nei, M., 1987. Molecular evolutionary genetics. Columbia University Press, New York, USA.

695 Olson, R.R., McPherson, R., 1987. Potential vs. realized larval dispersal: fish predation on
696 larvae of the ascidian *Lissoclinum patella* (Gottschaldt). J. Exp. Mar. Biol. Ecol. 110, 245-
697 256.

698 Palumbi, S.R., 1994. Genetic divergence, reproductive isolation, and marine speciation. Ann.
699 Rev. Ecol. Syst. 25, 547-572.

700 Palumbi, S.R., 2004. Marine reserves and ocean neighborhoods: the spatial scale of marine
701 populations and their management. Annu. Rev. Environ. Resour. 29, 31-68

702 Palumbi, S.R., Grabowsky, G., Duda, T., Geyer, L., Tachino, N., 1997. Speciation and
703 population genetic structure in tropical Pacific sea urchins. Evolution 51, 1506-1517.

704 Pancer, Z., Gershon, H., Rinkevich, B., 1994. Direct typing of polymorphic microsatellites in
705 the colonial tunicate *Botryllus schlosseri* (Ascidacea). Biochem. Bioph. Res. Co. 203, 646-
706 651.

707 Pascual, M., Balanyà, J., Latorre, A., Serra, L., 1997. Analysis of the variability of
708 *Drosophila azteca* and *Drosophila athabasca* populations revealed by random amplified
709 polymorphic DNA. J. Zool. Syst. Evol. Res. 35, 159-164.

710 Pérez-Portela, R., 2006. Phylogeny, population genetics and biology of the genus
711 *Pycnoclavella* (Ascidacea). PhD Thesis, University of Barcelona, Barcelona (Spain).

712 Pérez-Portela, R., Duran, S., Estoup, A., Turon, X., 2006. Polymorphic microsatellite loci
713 isolated from the Atlanto-Mediterranean colonial ascidian *Pycnoclavella* sp. (Ascidacea,
714 Tunicata). Mol. Ecol. Notes 6, 519-520.

715 Pérez-Portela, R., Duran, S., Palacín, C., Turon, X., 2007. The genus *Pycnoclavella*
716 (Ascidacea) in the Atlanto-Mediterranean region: a combined molecular and morphological
717 approach. Invertebr. Syst. (in press).

718 Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution.
719 Bioinformatics 14, 817-818.

720 Ramos-Onsins, S., Rozas, J., 2002. Statistical properties of a new neutrality test against
721 population growth. Mol. Biol. Evol. 19, 2092-2100.

722 Raymond, M., Rousset, F., 1995. GENEPOP (Version 1.2): population genetics software for
723 exact tests and ecumenicism. J. Hered. 86, 248-249.

724 Rinkevich, B., 2005. Natural chimerism in colonial urochordates. J. Exp. Mar. Biol. Ecol.
725 322, 93-109.

726 Roman, J., Palumbi, S.R., 2004. A global invader at home: population structure of the green
727 crab, *Carcinus maenas*, in Europe. Mol. Ecol. 13, 2891-2898.

728 Rousset, F., 1997. Genetic differentiation and estimation of gene flow from F-statistics under
729 isolation by distance. Genetics 145, 1219-1228.

- 730 Rozas, J., Sánchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA
731 polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496-2497.
- 732 Schneider, S., Roessli, D., Excoffier, L., 2000. ARLEQUIN ver 2000: A software for
733 population genetics data analysis, University of Geneva, Switzerland.
- 734 Shanks, A.L., Grantham, B.A., Carr, M.H., 2003. Propagule dispersal distance and the size
735 and spacing of marine reserves. *Ecol. Appl.* 13, 159-169.
- 736 Sites, J.W., Marshall, J.C., 2003. Delimiting species: a Renaissance issue in systematic
737 biology. *Trends Ecol. Evol.* 18(9), 462-470.
- 738 Slatkin, M., 1987. Gene flow and the geographic structure of natural populations. *Science*
739 236, 787-792.
- 740 Sommerfeldt, A.D., Bishop, J.D.D., Wood, C.A., 2003. Chimerism following fusion in a
741 clonal ascidian (Urochordata). *Biol. J. Linn. Soc.* 79, 183-192.
- 742 Stoner, D.S., Quattro, J.M., Weissman, I.L., 1997. High polymorphic microsatellite loci in the
743 colonial ascidian *Botryllus schlosseri*. *Mol. Mar. Biol. Biotech.* 6, 163-171.
- 744 Stoner, D.S., Ben-Shlomo, R., Rinkevich, B., Weissman, I.L., 2002. Genetic variability of
745 *Botryllus schlosseri* invasions to the east and west coasts of the USA. *Mar. Ecol. Prog. Ser.*
746 243, 93-100.
- 747 Svane, I., Young, C.M., 1989. The ecology and behaviour of ascidian larvae. *Oceanogr. Mar.*
748 *Biol.* 27, 45-90.
- 749 Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA
750 polymorphism. *Genetics* 123, 585-595.
- 751 Tarjuelo, I., Posada, D., Crandall, K.A., Pascual, M., Turon, X., 2001. Cryptic species of
752 *Clavelina* (Ascidiacea) in two different habitats: harbours and rocky littoral zones in the
753 northwestern Mediterranean. *Mar. Biol.* 139, 455-462.
- 754 Tarjuelo, I., Posada, D., Crandall, K.A., Pascual, M., 2004. Phylogeography and speciation of
755 colour morphs in the colonial ascidian *Pseudodistoma crucigaster*. *Mol. Ecol.* 13, 3125-3136.

756 Templeton, A.R., 2001. Using phylogeographic analyses of gene trees to test species status
757 and processes. *Mol. Ecol.* 10, 779-791.

758 Tintoré, J., La Violette, P.E., Blade, I., Cruzado, A., 1988. A study of an intense density front
759 in the Eastern Alboran Sea: the Almeria-Oran front. *J. Phys. Oceanogr.* 18, 1384-1397.

760 Turon, X., Tarjuelo, I., Duran, S., Pascual, M., 2003. Characterising invasion processes with
761 genetic data: an Atlantic clade of *Clavelina lepadiformis* (Ascidacea) introduced into
762 Mediterranean harbours. *Hydrobiologia* 503, 29-35.

763 Waples, R.S., Gaggioti, O., 2006. What is a population? An empirical evaluation of some
764 genetic methods for identifying the number of gene pools and their degree of connectivity.
765 *Mol. Ecol.* 15, 1419-1439.

766

767 Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population
768 structure. *Evolution* 38, 1358-1370.

769 Whalan, S., Johnson, M.S., Harvey, E., Battershill, C., 2005. Mode of reproduction,
770 recruitment, and genetic subdivision in the brooding sponge *Haliclona* sp. *Mar Biol* 146, 425-
771 433.

772 Yund, P.O., O'Neil, P.G., 2000. Microgeographic genetic differentiation in a colonial ascidian
773 (*Botryllus schlosseri*) population. *Mar. Biol.* 137, 583-588.

774 Zane, L., Bargelloni, L., Patarnello, T., 2002. Strategies for isolation microsatellite: a review.
775 *Mol. Ecol.* 11, 1-6.

Table 1. Diversity measures of each population of *Pycnoclavella communis* studied. Sample size (*n*), number of polymorphic sites (Np), number of haplotypes (Nh), nucleotide diversity (π) and haplotype diversity (*h*) are indicated. Standard deviations are given for π and *h*.

Population	<i>n</i>	Np	Nh	π	<i>h</i>
Carboneras (CA)	18	4	4	0.0021± 0.00059	0.542± 0.123
Aguamarga (AG)	15	1	2	0.00044± 0.00023	0.248± 0.131
Cabo de Palos (CP)	13	2	3	0.00054± 0.00030	0.295± 0.156
Lineage B	46	8	7	0.00346± 0.00041	0.648± 0.061
Pta de la Mona (PM)	19	8	4	0.0027± 0.00097	0.509± 0.117
Cerro Gordo (CG)	19	33	3	0.01165± 0.00406	0.433± 0.117
Blanes (BL)	16	7	6	0.0028± 0.00073	0.683± 0.120
Tossa (T)	16	4	5	0.00233± 0.00041	0.75± 0.078
Palamós (P)	19	8	6	0.0046± 0.00051	0.678± 0.093
Lineage A	89	43	22	0.00982 ± 0.00099	0.858 ± 0.0006
Atlantic sites (ATL)	3	26	3	0.03112± 0.0116	1± 0.272
Total	138	92	32	0.04451± 0.00194	0.903± 0.013

Table 2. Neutrality tests on Cytochrome c oxidase data inferred from the Maximum Likelihood method for lineage B, for the Mediterranean clade within lineage A and for each population of *Pycnoclavella communis*. * $P < 0.05$; NS not significant.

Populations	Tajima's D-test	Fu's Fs test	R ₂ test
Lineage B	0.144 NS	0.066 NS	0.1171 NS
Aguamarga	-0.512 NS	1.187 NS	0.1238 NS
Cabo de Palos	-1.468 NS	-1.401 NS	0.1804 NS
Carboneras	0.591 NS	0.654 NS	0.1745 NS
Lineage A-Mediterranean clade	-1.833280*	3.74159 NS	0.0632 NS
Pta de la Mona	-1.65514*	0.32339 NS	0.1449 NS
Cerro Gordo	-1.22666 NS	9.33573 NS	0.1565 NS
Blanes	-0.7536 NS	-1.19943 NS	0.1073 NS
Tossa	0.26698 NS	-0.75176 NS	0.1598 NS
Palamós	0.49457 NS	0.39627 NS	0.1510 NS

Table 3. F_{st} values of pairwise comparisons between the five Mediterranean populations within lineage A of *Pycnoclavella communis* obtained from the Cytochrome c oxidase data (see abbreviations of the populations in Material and Methods)

	PM	CG	BL	T
PM				
CG	0.05059			
BL	0.40771**	0.44776**		
T	0.3755**	0.41556**	0.28333**	
P	0.40643**	0.44444**	0.31924**	0.07292*

* Significant at $P < 0.05$ and ** highly significant at $P < 0.01$ after 10,000 permutations of haplotypes between locations.

Table 4. Analysis of the molecular variance (AMOVA) of *Pycnoclavella communis* from Cytochrome c oxidase data. AMOVA of Mediterranean populations within lineage A grouped in South and North areas and for the whole Mediterranean area without grouping.

Source of variation	d.f.	Sum of squares	Variance components	% of variation	P	Fixation indices
Among groups	1	7.590	0.13630 Va	26.61	0.01	F_{ct} : 0.26614
Among populations within groups	3	4.834	0.07431 Vb	14.51*	< 0.001	F_{sc} : 0.19771
Within populations	84	25.329	0.30154 Vc	58.88*	< 0.001	F_{st} : 0.41123
Total	88	37.735	0.51215			
AMOVA without grouping						
Among populations	4	12.424	0.15782 Va	34.36*	< 0.001	F_{st} : 0.34357
Within populations	84	25.329	0.30154 Vb	65.64		
Total	88	37.753	0.45936			

Table 5. Summary of genetic variation for eight microsatellite loci at four Mediterranean populations of *Pycnoclavella communis* within lineage A. Number of alleles, number of observed heterozygotes (Ho), number of expected heterozygotes (He), inbreeding coefficient (F_{is}) and number of individuals amplified excluding failed amplifications (N) are given (see abbreviations of the populations in Material and Methods section)

LOCUS	POPULATION			
	CG	PM	BL	T
PYC-1				
n° alleles	monomorphic	2	monomorphic	monomorphic
Ho		2		
He		9.081080		
F_{is}		0.784***		
N	30	19	8	27
PYC-9				
n° alleles	2	2	2	4
Ho	1	3	1	6
He	4.789474	4.354839	1	6.758621
F_{is}	0.800*	0.318	0.000	0.116
N	10	16	7	15
PYC-35				
n° alleles	3	2	3	null
Ho	9	5	2	-
He	7.693877	3.666667	6.757576	-
F_{is}	-0.174	-0.400	0.710**	-
N	25	8	17	0
PYC-50				
n° alleles	5	null	null	7
Ho	6	-	-	5
He	15.929822	-	-	20.905659
F_{is}	0.627***	-	-	0.764***
N	29	0	0	27
PYC-53				
n° alleles	3	2	5	3
Ho	4	0	2	6
He	17	1.818182	12.744681	9.803922
F_{is}	0.768***	1	0.846***	0.393*
N	28	6	24	26
PYC-67				
n° alleles	9	8	3	monomorphic
Ho	12	14	6	
He	22.894735	15.540541	13.283019	
F_{is}	0.480***	0.102	0.553***	
N	29	19	27	4
PYC-89				
n° alleles	5	monomorphic	3	3
Ho	8		2	13
He	13.030304		3.384616	9.413794
F_{is}	0.393**		0.429	-0.400
N	17	13	7	15
PYC-92				
n° alleles	4	null	3	null
Ho	7	-	15	-
He	7	-	11.487805	-
F_{is}	0.000	-	-0.316*	-
N	12	0	21	0
Mean n° alleles/ location	4	2.83	2.85	3.1
F_{is}	0.449***	0.248**	0.405***	0.300**
Total HWE	***	**	***	***

Significant values are indicated as * for $P < 0.05$, ** for $P < 0.01$ and *** $P < 0.001$

Table 6. Percentage of failed amplifications for each locus and each Mediterranean population within Lineage A of *Pycnoclavella communis* (see abbreviations of the populations in Material and Methods section)

Locus	% failed amplification
PYC-1	24.5%
PYC-9	56.4%
PYC-35	54.5%
PYC-50	49%
PYC-53	23.6%
PYC-67	28.2%
PYC-89	52.7%
PYC-92	70%
Population	
CG	25%
PM	50%
BL	53.7%
T	52.5%
Total	45%

Table 7. F_{st} (below the diagonal) and Rho_{st} (above the diagonal) values of pairwise comparisons between the four Mediterranean populations within lineage A of *Pycnoclavella communis* obtained from microsatellite data (see abbreviations of the populations in Material and Methods). All values were significant at $P < 0.05$.

	PM	CG	BL	T
PM		0.494	0.789	0.592
CG	0.359		0.201	0.212
BL	0.352	0.367		0.538
T	0.396	0.319	0.386	

Table 8. Global single-locus F_{st} and Rho_{st} values of pairwise comparisons for the microsatellites analysed in *Pycnoclavella communis*.

	Locus							
	PYC-1	PYC-9	PYC-35	PYC-50	PYC-53	PYC-67	PYC-89	PYC-92
F_{st}	0.563	0.145	0.450	0.532	0.308	0.359	0.149	0.363
Rho_{st}	0.636	0.026	0.470	0.303	0.747	0.018	0.546	0.381

Table 9. Analysis of the molecular variance (AMOVA) for Mediterranean populations within lineage A of *Pycnoclavella communis* from microsatellite data. AMOVA of Mediterranean populations grouped in South and North areas and for the whole Mediterranean area without grouping.

Source of variation	d.f.	Sum of squares	Variance components	% of variation	P	Fixation indices
Among groups	1	53.818	-0.07102 Va	-2.51	0.665	F_{ct} : -0.02512
Among populations within groups	2	119.279	1.08095 Vb	38.23*	0.000	F_{sc} : 0.37290
Within populations	214	389.022	1.81786 Vc	64.29*	0.000	F_{st} : 0.35715
Total	217	562.119	2.82779			
AMOVA without grouping						
Among populations	3	173.098	1.03355 Va	36.25*	0.000	F_{st} : 0.36247
Within populations	214	389.022	1.81786 Vb	63.75		
Total	217	562.119	2.85141

Figure 1. Map of the localities sampled in the Mediterranean Sea.

Figure 2. Maximum Likelihood tree of the obtained haplotypes, estimated using the model of evolution TVM+I+G. Bootstrap values for the branches with more than 50% support are shown.

Figure 3. Median-joining network. Simple lines represent one mutation step between haplotypes, roman numerals are indicating the number of mutations when there are more than one. Red circles indicate missing haplotypes. Haplotypes of *Pycnoclavella communis* are named with arabic numbers. The area of the circles is proportional to the frequency of each haplotype and colours are indicating different populations. * Non-synonymous substitutions.

Appendix I. Relative haplotype frequencies for each population of *Pycnoclavella communis* studied. Private haplotypes are shown in bold.

	Carboneras	Aguamarga	Cabo de Palos	Pta de la Mona	Cerro Gordo	Blanes	Tossa	Palamós	Atlantic sites	TOTAL	Acc number
H1	0.167	-	-	-	-	-	-	-	-	0.0217	Pending
H2	0.667	0.867	-	-	-	-	-	-	-	0.181	Pending
H3	0.111	0.133	-	-	-	-	-	-	-	0.0289	Pending
H4	0.0556	-	-	-	-	-	-	-	-	0.0072	Pending
H5	-	-	-	0.211	-	-	-	-	-	0.0289	Pending
H6	-	-	-	0.684	0.737	-	-	-	-	0.195	Pending
H7	-	-	-	0.0526	-	-	-	-	-	0.0072	Pending
H8	-	-	-	0.0526	-	-	-	-	-	0.0072	Pending
H9	-	-	-	-	0.211	-	-	-	-	0.0289	Pending
H10	-	-	-	-	0.0526	-	-	-	-	0.0072	Pending
H11	-	-	0.846	-	-	-	-	-	-	0.0797	Pending
H12	-	-	0.0769	-	-	-	-	-	-	0.0072	Pending
H13	-	-	0.0769	-	-	-	-	-	-	0.0072	Pending
H14	-	-	-	-	-	0.562	-	-	-	0.0652	Pending
H15	-	-	-	-	-	0.125	-	-	-	0.0145	Pending
H16	-	-	-	-	-	0.125	-	-	-	0.0145	Pending
H17	-	-	-	-	-	0.0625	-	-	-	0.0072	Pending
H18	-	-	-	-	-	0.0625	-	-	-	0.0072	Pending
H19	-	-	-	-	-	0.0625	-	-	-	0.0072	Pending
H20	-	-	-	-	-	-	0.438	0.526	-	0.123	Pending
H21	-	-	-	-	-	-	0.25	-	-	0.0289	Pending
H22	-	-	-	-	-	-	0.188	-	-	0.0217	Pending
H23	-	-	-	-	-	-	0.0625	-	-	0.0072	Pending
H24	-	-	-	-	-	-	0.0625	-	-	0.0072	Pending
H25	-	-	-	-	-	-	-	0.0526	-	0.0072	Pending
H26	-	-	-	-	-	-	-	0.263	-	0.0362	Pending
H27	-	-	-	-	-	-	-	0.0526	-	0.0072	Pending
H28	-	-	-	-	-	-	-	0.0526	-	0.0072	Pending
H29	-	-	-	-	-	-	-	0.0526	-	0.0072	Pending
H30	-	-	-	-	-	-	-	-	0.25	0.0072	AM403667
H31	-	-	-	-	-	-	-	-	0.5	0.0072	AM403666
H32	-	-	-	-	-	-	-	-	0.25	0.0072	AM403665

Appendix II. Allele frequencies for eight microsatellite loci from 4 Mediterranean populations of *Pycnoclavella communis*. Private alleles are shown in bold.

		Location			
Locus	Allele	CG	PM	BL	T
PYC-1	158	0	0.632	0	0
	178	1	0.368	1	1
PYC-9	88	0	0	0	0.133
	102	0.650	0.844	0.929	0.733
	136	0.350	0.156	0.071	0.067
	142	0	0	0	0.067
PYC-35	215	0.160	0	0	-
	225	0.020	0	0	-
	280	0	0	0.088	-
	282	0	0	0.765	-
	286	0	0	0.147	-
	293	0.820	0	0	-
	321	0	0.688	0	-
	329	0	0.312	0	-
PYC-50	202	0	-	-	0.111
	220	0	-	-	0.204
	232	0	-	-	0.111
	236	0.621	-	-	0.148
	238	0	-	-	0.389
	240	0.259	-	-	0
	242	0.017	-	-	0.019
	244	0.017	-	-	0
	246	0.086	-	-	0.019
PYC-53	258	0	0.167	0	0.038
	266	0.339	0.833	0	0
	270	0.518	0	0	0
	278	0	0	0.042	0
	280	0	0	0.104	0
	282	0.143	0	0.667	0.769
	286	0	0	0.146	0.192
	292	0	0	0.042	0
PYC-67	265	0.155	0.053	0.259	0
	273	0.086	0.289	0	0
	275	0.155	0	0	0
	277	0.086	0.079	0.074	0
	279	0	0.053	0	0
	281	0	0.079	0.667	1
	283	0.397	0.289	0	0
	289	0.017	0.132	0	0
	298	0.034	0	0	0
	301	0.034	0.026	0	0
	303	0.034	0	0	0
PYC-89	171	0	0	0.143	0
	185	0.147	0	0.143	0
	191	0.324	1	0.714	0.133
	215	0.324	0	0	0.433
	223	0.147	0	0	0.433
	227	0.059	0	0	0
PYC-92	115	0.125	-	0.405	-
	127	0.000	-	0.548	-
	129	0.625	-	0.048	-
	133	0.083	-	0	-
	157	0.167	-	0	-