**Mixed but not admixed: A spatial analysis of genetic variation of an invasive
ascidian on natural and artificial substrates**

Víctor Ordóñez¹, Marta Pascual¹, Marc Rius^{2,3} and Xavier Turon³

¹ Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avgda.
Diagonal 645, edifici annex, 08028, Barcelona, Spain.

² Department of Evolution and Ecology, University of California, One Shields Avenue,
Davis, California 95616, United States of America.

³ Centre d'Estudis Avançats de Blanes (CEAB-CSIC), Accés a la Cala S. Francesc 14,
17300 Blanes (Girona), Spain.

Corresponding author: Xavier Turon

E-mail: xturon@ceab.csic.es

Fax: +34972337806

Phone: +34972336101

ABSTRACT

Following the introduction to a new area (pre-border dispersal), post-border processes determine the success in the establishment of non-indigenous species (NIS). However, little is known on how these post-border processes shape the genetic composition of NIS at regional scales. The aim of the present work is to analyse genetic variation in introduced populations along impacted coastlines as a tool to infer post-border dynamics. We used as a model system the ascidian species *Microcosmus squamiger* that has been introduced worldwide. This species can colonize and grow fast on man-made artificial structures, impacting activities such as mariculture. However it can also establish itself in natural substrates, thus altering natural communities and becoming an ecological problem. We genotyped 302 individuals from eight populations established on natural and artificial substrates in the northwestern Mediterranean Sea, using six microsatellite loci. We then compared the resulting genotypes with those found within the native range of the species. We found high levels of genetic diversity and allelic richness in all populations, with an overall deficit of heterozygotes. Autocorrelation analyses showed that there was no within-population genetic structure (at a scale of tens of meters); likewise, no significant differentiation in pairwise comparisons between populations (tens of kilometres apart) and no isolation-by-distance pattern was found. The results suggest that *M. squamiger* has a natural capacity for high dispersal from one patch of hard substrate to another and no differences whatsoever could be substantiated between natural and artificial substrates. Interestingly, two groups of genetically differentiated individuals were detected that were associated with the two ancestral source areas of the worldwide expansion of the species. Individual assignment tests

showed the coexistence of individuals of these two clusters in all populations but with little interbreeding among them as the frequency of admixed individuals was only 15%. The mechanism responsible for the maintenance of these different genetic pools is unknown, but it apparently does not compromise the colonization potential of the introduced populations.

INTRODUCTION

The artificial introduction of species to new areas, defined as pre-border (Forrest et al. 2009) or extra-range dispersal (Wilson et al. 2009) is mediated by several vectors (Carlton and Geller 1993; Ruiz et al. 1997; Wonham et al. 2001) and is an on-going process that will continue in the foreseeable future despite prevention efforts (Bax et al. 2001; Hulme 2006). However, the establishment of a species translocated to a new area depends on the success of post-border processes, i.e., those involving the establishment and spread after the initial inoculation (Forrest et al. 2009). Such processes determine whether the species will become invasive or not, and open valuable opportunities for management and prevention of invasive species (Wotton and Hewitt 2004; Forrest et al. 2009; Airolidi and Bulleri 2011). Survival, reproduction, dispersal, and local environmental conditions are the main factors that determine the success of an introduced population during the post-border stage (Blackburn et al. 2011). A key process is the regional spread of the species that, contrary to what happens in extra-range dispersal, relies heavily on its natural dispersal abilities, as well as on human-mediated transport at local scales (Wasson et al. 2001; Branch and Steffani 2004).

A common feature of many successful invaders is that their colonization histories often involve multiple introductions from multiple sources (Sakai et al. 2001;

Geller et al. 2010), which in the sea is often accomplished through repeated transport following the main routes of ship traffic or via aquaculture activities (Roman and Darling 2007). This often leads to a high genetic diversity of introduced populations through admixture of genetically differentiated sources (Geller et al. 2010). Furthermore, the interaction between different genetic pools (e.g., McIvor et al. 2001; Meusnier et al. 2002; Kelly et al. 2006; Pineda et al. 2011) from different origins and, potentially, different ecophysiological characteristics, can significantly affect the success of introduced populations in a way that remains largely unexplored.

In heavily urbanized coastlines, dispersal is also favoured by the existence of marinas and harbours, with associated shipping activities. While primary introductions can depend on ship traffic between main ports, a network of marinas with boating activities can contribute to secondary dispersal (Lambert and Lambert 1998; Wasson et al. 2001; Lacoursière-Roussel et al. 2012). Natural dispersal between suitable habitats can lead to stepping-stone models of dispersal, while boat-mediated dispersal can break down relationships between geographical distance and connectivity at local scales (Lacoursière-Roussel et al. 2012).

Artificial substrates can act as dispersal corridors for exotic rocky-bottom species (Bulleri and Aioldi 2005; Glasby et al. 2007; Vaselli et al. 2008; Bulleri and Chapman 2010; Aioldi and Bulleri 2011) and can represent a source of infiltration to natural benthos by non-indigenous species (Simkanin et al. 2012). These structures do not function as surrogates from natural substrates, as structure and dynamics of communities on artificial hard substrate are different from those in adjacent natural rocky bottoms (e.g., Pinn et al. 2005; Tyrrell and Byers 2007; Bulleri and Chapman 2010; Aioldi and Bulleri 2011). In general, species richness is lower in communities on artificial substrate than on well-established natural communities (e.g., Bacchiocchi and

Airolidi 2003), and many ecological processes such as recruitment, competition or predation may differ between natural and artificial habitats (Bulleri and Chapman 2010; Dumont et al. 2011). Overall, natural communities are expected to be able to influence propagule settlement more than artificial assemblages and, thus, to be less prone to invasion (Stachowicz et al. 2005; Tyrrell and Byers 2007; Dumont et al. 2011, but see Osman and Whitlatch 2007). Therefore, in order to understand patterns of colonization and connectivity with possible source areas (e.g. Fauvelot et al. 2009), it is essential to compare the genetic composition and diversity of introduced populations on natural and artificial substrates.

Among marine organisms, ascidians have become a worldwide problem as invader species (Lambert and Lambert 1998, 2003; Dijkstra et al. 2007; Lambert 2007; Locke 2009). Non-indigenous ascidians rapidly colonize artificial substrates in harbours such as pilings, docks, floating pontoons, boat hulls and buoys (Lambert 2002; Lambert and Lambert 2003). From these artificial substrates some species can colonize adjacent natural environments (Lambert 2007; Rius et al. 2009). Ascidians expand their ranges naturally via their non-feeding larvae, which only allow a restricted dispersal. This group is therefore a good model to study patterns of dispersal and connectivity of introduced species at small scales. Pre-border processes in non-indigenous ascidian species have been studied in a number of species with the aid of genetic tools, focusing on relatedness among colonized areas and on inferring the putative origin of colonizers (e.g. López-Legentil et al. 2006; Barros et al. 2009; Zhan et al. 2010; Goldstien et al. 2011; Lejeusne et al. 2011; Pineda et al. 2011; Rius et al. 2012). However, post-border dispersal processes are less studied in ascidians (e.g. *Styela clava*, Goldstien et al. 2010; *Ciona intestinalis*, Zhan et al. 2012; *Perophora japonica*, Pérez-Portela et al. 2012).

The solitary ascidian *Microcosmus squamiger* Michaelson 1927, is native to Australia (Michaelson 1927; Kott 1985; Rius et al. 2008a) and has been introduced in many parts of the world, such as the eastern Atlantic, western Mediterranean, California, South Africa, India and New Zealand (Rius et al. 2012 and references therein). *M. squamiger* has been introduced as a stowaway in ports and marinas via ballast water or ship fouling (Rius et al. 2008a) and has the capacity to colonize natural substrates (Turon et al. 2007) where it forms dense populations (Rius et al. 2009), thus this species represents a valuable model to study behaviour and dispersal capacities of invasive populations. Genetic data showed that its colonization process involved two source areas genetically differentiated (eastern and western Australia, Rius et al. 2012) which contributed to the genetic makeup of the introduced populations. Pre-border processes (global genetic patterns) in this species have been studied by Rius et al. (2008a, 2012), but there is no information on fine-scale genetic patterns and how it spreads at local scales (i.e., at the post-border level).

M. squamiger can entirely carpet artificial habitats (Turon et al. 2007; Rius et al. 2009). A plausible consequence of a fast colonization of coastal infrastructures would be a low genetic diversity of these populations, as found for other benthic invertebrates (Fauvelot et al. 2009), with high kinship values within the aggregations. Natural habitats have less dense populations of *M. squamiger* (authors' pers. obs.) and interspecies competition could be exerting some control over them, impeding monopolization of substrate and allowing multiple introductions from diverse sources over time, resulting in higher genetic diversity.

The study of genetic variability and population connectivity of invasive species can unravel post-border dispersal dynamics and is a necessary step to design management and intervention plans (Forrest et al. 2009). In this sense, the main goal of

the present study was to assess post-border, fine-scale dispersal in artificial and natural substrates of introduced populations of *M. squamiger* by analysing genetic structure within and between populations. Our aims were (i) to study the genetic variability of both natural and artificial substrate populations, (ii) to determine if fine-scale population structure exists within populations and varies according to the type of substrate, (iii) to assess inter-population connectivity related to substrate type and geographical distance and, (iv) to analyse the degree of admixture at the individual and population levels with respect to the two recognized genetic pools in the native area of the species.

MATERIAL AND METHODS

Sampling area

We sampled eight sites along a 50 kilometer (km) stretch of highly urbanized coast in the northwestern Mediterranean Sea (Fig. 1). There were two big commercial ports north and south of the studied coastline, the port of Barcelona located 30 km north-eastwards, and the Port of Tarragona 22 km south-westwards. These ports can act as entry points for the species in the area. This coast mainly consists of sandy areas interspersed with natural rocky outcrops and artificial breakwaters (spaced at most ca. 3 km). It provides, therefore, an adequate setting for testing stepping-stone dispersal of the studied species, which does not settle on soft substrate. All sampled sites were located in the open coastal zones, that is, they are located outside confined areas such as marinas or embayments. Four of the sites were found on natural substrates: Garraf-natural (GAN), Sitges-natural (SIN), El Roc de Sant Gaietà (SG), Torredembarra (TB); and four on artificial substrates: Garraf-artificial (GAA), Sitges-artificial (SIA),

Cubelles (CU) and Segur de Calafell (SC) (Fig. 1, Table S1). Samples were collected in 2010 from natural rocky walls or groynes and breakwaters at depths between 1 and 4 m through snorkelling or SCUBA. Overall, we studied 302 individuals; 156 collected from natural substrate and 146 from artificial substrate (Table S1).

In order to investigate fine-scale (within population) genetic structure, we collected samples along a fixed 90 m horizontal transect marked with a measuring tape in two of the natural substrate sites (Sitges and El Roc de Sant Gaietà) and two with artificial substrate (Cubelles and Segur de Calafell). Individuals were collected at 7 points along these transects at increasing distance from the initial point (Table S2). At each point all individuals situated within a 20 cm radius were collected (resulting in sample sizes of 4-15 individuals per point). The design allowed for comparison of relatedness among individuals located 0 to 90 m apart.

The remaining localities were sampled by collecting individuals haphazardly (i.e. picking individuals randomly but at distances larger than 0.5 m) across a total linear distance of ca. 100 m. These and the previous samples were used to test connectivity as related to substrate type and geographic distance. We obtained thus the two following datasets: one to assess genetic variability and differentiation among all populations; and the other to assess fine-scale genetic structure in the four populations sampled following the transect line. Once in the laboratory, all individuals were dissected and a piece of muscular tissue preserved in absolute ethanol and stored at -80°C.

DNA extraction and amplification

We used six microsatellite loci isolated from this species (Rius et al. 2008b): MS6, MS7, MS10, MS11, MS12, and MS13. These microsatellites do not show linkage disequilibrium so they can be considered independent loci (Rius et al. 2008b). The

genomic DNA extraction was performed using the *Squishing Buffer* (SB) + *proteinase K* method, originally developed for *Drosophila* (Gloor et al. 1993).

The PCR amplification was performed with a final volume of 20µl. MS7 and MS13 were amplified together at 57°C annealing temperature, with 4µl of buffer 5X (Green GoTaq, Promega), 2.5µl of dNTPs (1mM), 1.8µl of MgCl₂ (25mM), 8.9µl of H₂O, 0.6µl of DMSO, 0.2µl of Taq corresponding to 1 unit (GoTaq, Promega), 0.25µl of each primer (10µM) and 1µl of DNA. The other microsatellites were amplified separately using 4µl of buffer 5X (Green GoTaq, Promega), 2.5µl of dNTPs (1mM), 3µl of MgCl₂ (25mM), 7.7µl of H₂O, 0.6µl of DMSO, 0.2µl of Taq corresponding to 1 unit (GoTaq, Promega), 0.25µl of each primer (10µM) and 1µl of DNA. The annealing temperature was 50°C for MS10, 53°C for MS6 and 57°C for MS11 and MS12. The PCRs started with an initial denaturation at 94°C for 5 minutes, followed by 35 cycles of a denaturation step at 94°C for 1 minute, an annealing step at the corresponding temperature for 30 seconds and an elongation step at 72°C for 30 seconds, with a final extension at 72°C for 5 minutes. The forward primer of each locus was marked with an appropriate fluorochrome (Rius et al. 2008b). We used the GeneMapper® software (version 3.7, Applied Biosystems, 2004) to assign allele sizes to each microsatellite and to genotype each individual.

We had a small number of failed amplifications for three of the loci. This was despite repeated attempts using different PCR conditions and new DNA extractions. The percentage of failed amplifications was low and varied among loci (0.3% for MS6; 3.64% for MS10; 0.66% for MS12).

We also performed comparisons of the microsatellite dataset here obtained with the genotypes obtained by Rius et al. (2012) from two native localities (Manly and Bunbury). These samples were genotyped in the same machine but at different years

than the present study, and the original chromatograms of these samples were manually checked for possible bias on allele calling between studies. No correction was necessary.

Analysis of genetic variability

The program Microsatellite toolkit v. 3.1.1 (Park 2001) was used to transform the data files into the adequate formats for the different programs used. In order to compare genetic diversity patterns among populations with different numbers of individuals we calculated the allelic richness corrected per sample size with the program Fstat v. 2.9.3.2 (Goudet 2002). We obtained the number of alleles and both the expected and observed heterozygosities for each locus and population using GeneClass v. 2 (Piry et al. 2004). Deviations from Hardy-Weinberg Equilibrium were tested with the inbreeding coefficient F_{IS} , and its significance assessed by 10,000 bootstrap replicates with the program Genetix v. 4.05.2 (Belkhir et al. 2004). In order to compare genetic variability between populations in natural and artificial substrate, the permutation test implemented in Fstat was used.

Within population spatial genetic structure

In order to assess fine-scale genetic structure, we used autocorrelation analysis, which allows an assessment of the scale at which discontinuities occur (Heywood 1991). We designed the spatial sampling across transects to evaluate potential genetic consequences of dispersal over fine spatial scales (Loiselle et al. 1995). Given all possible distances between the different sampling points (Table S2), we established 18 distance classes (in meters): 0.2, 2.5, 5, 7.5, 10, 15, 17.5, 20, 30, 35, 37.5, 40, 50, 70, 80, 85, 87.5 and 90. These distances designate the end-points of each distance-class. Distances were chosen to obtain the highest number of intervals but with enough data

points at each interval to have at least 20 pairwise comparisons within each class (overall mean: 56.25 comparisons per distance class). The class 0.2 refers to individuals collected within each transect point and separated by less than 20 cm.

For individual comparisons we computed two statistics: the kinship coefficient (Loiselle et al. 1995) and the $\hat{a}$ index defined by Rousset (2000). The kinship coefficient is a similarity measure that estimates the probability that two alleles of an autosomal locus are identical by descent (Ritland 1996). The advantage of this index is that it is not dependent on Hardy-Weinberg equilibrium conditions (Hardy and Vekemans 1999; Fenster et al. 2003). The $\hat{a}$ index is a distance measure analogous to the $F_{ST}/(1-F_{ST})$ ratio (see Rousset 2000 for details). The 95% confidence intervals of no association at a given distance interval were assessed through 1,000 randomizations, permuting the position of the individuals along the transect, which effectively eliminated all spatial structure but preserved sample sizes and individual genotypes. We performed these analyses using the program SPAGeDi v. 1.2 (Hardy and Vekemans 2002). In order to compare the kinship coefficient and the $\hat{a}$ index between natural and artificial substrates, SigmaStat v. 3.11 (Systat Software, Inc.) was used to perform the non-parametric Mann-Whitney U -test.

Genetic differentiation among populations

Recent studies have questioned the suitability of commonly used G_{ST} - and F_{ST} -like estimators to assess population differentiation, as they are highly dependent on the variability of the marker used, and new estimators have been proposed to overcome these problems (Hedrick 2005; Jost 2008). Nevertheless, the issue is still in debate (Whitlock 2011), and it seems advisable for the time being to use both traditional and new estimators (Meirmans and Hedrick 2011). We have therefore assessed population

genetic differentiation using conventional F_{ST} estimates and the new estimator D_{est} as in equation 12 of Jost (2008). The former was calculated with Genetix using the estimator of Weir & Cockerham (1984). The significance of the pairwise values was assessed with 1,000 permutations of data, after correcting for multiple comparisons using the Benjamini-Yekutieli method described in Narum (2006). For D_{est} estimates we used the R package DEMETics v. 0.8.1 (Gerlach et al. 2010). We calculated a confidence interval around the obtained values by 1,000 bootstrap replicates. In this procedure, alleles are automatically randomized over populations when the compared populations are in Hardy-Weinberg equilibrium; otherwise, genotypes are randomized. The programme automatically adjusts confidence intervals for multiple comparisons. Significant differentiation was inferred when this confidence interval excluded zero. The program GenePop On The Web v. 4.0.10 (Raymond and Rousset 1995) was used to test whether there is isolation by distance among the studied populations comparing the D_{est} and F_{ST} pairwise matrices with the geographic distance matrix using a Mantel test. Analyses of molecular variance (AMOVA) grouping populations by substrate (natural and artificial) were performed using a F_{ST} -like statistic with the program Arlequin v. 3.5.1 (Excoffier et al. 2005).

Admixture analysis

In order to assess whether there was any overall structure in our microsatellite dataset, we used the program STRUCTURE v. 2.3 to find the most likely number of genetically differentiated clusters (K) (Pritchard et al. 2000; Hubisz et al. 2009). We used the admixture and loc prior model because it performs better than other models for detecting genetic structure even in situations of low levels of genetic divergence or a limited number of loci (Hubisz et al. 2009). Twenty independent runs were performed

for increasing values of K from 1 to 8 (number of populations) using 500,000 iterations and a burn-in period of 50,000. We ran CLUMPP v. 1.1.2 (Jakobsson and Rosenberg 2007) to merge the results of the 20 runs for the most likely K and DISTRICT v.1.1 (Rosenberg 2004) was used to graphically display the results. Using the 20 runs we plotted both the IncK statistic (Evanno et al. 2005) and the log probability of the data (LnP(D)) as a function of K and looked for the value that best captured the structure of the data. Additionally, we used a K -means clustering algorithm available in the R package *adeigenet* 1.2-8 (Jombart 2008) to find groups of individuals that maximize between-group genetic variation (function *find.clusters*). We used the same K as obtained with STRUCTURE to compare the groups recovered by both methods.

In order to check for structure in our data with a different approach, we ran a factorial correspondence analysis (FCA) using Genetix. FCA provides a graphical representation of genetic distances among individuals without any *a priori* clustering, and also evaluates the relative contribution of each allele to the ordination found.

We used Discriminant Analysis of Principal Components (DAPC, Jombart et al. 2010), to further analyse our data. DAPC is a recently developed technique that extracts information from genetic datasets (multivariate in nature) by first performing a principal component analysis (PCA) on pre-defined groups or populations, and then using the PCA factors as variables for a discriminant analysis (DA). In order to assess the relationships of the groups defined with STRUCTURE with individuals from two populations, Bunbury (W Australia) and Manly (E Australia), representative of the two source areas of the worldwide expansion of the species (Rius et al. 2012), we ran DAPC (function *dapc* in the package *adeigenet*) with our samples and the Australian populations. Variables were centred but not scaled, and 30 principal components of PCA were retained and input to DA. Additionally, an AMOVA grouping individuals as

per genetic groups derived from STRUCTURE results was performed with Arlequin as explained above.

RESULTS

Genetic variability of natural and artificial substrate populations

There were large differences in the genetic variability of the different loci, ranging from 3 to 13 alleles. Mean total allelic richness was similar between populations in natural and artificial substrates (natural substrate mean $\pm$ SE = 24.530 ± 0.557 ; artificial substrate = 25.136 ± 1.829). No significant differences were found with a permutation-based method between natural and artificial populations (P -values for allelic richness = 0.802; H_o = 0.769; gene diversity = 0.486; F_{IS} = 0.688; 10,000 permutations). The inbreeding coefficient F_{IS} across loci (Table 1) was significant in one population from natural substrate (SIN) and three populations from artificial substrate (GAA, CU and SC). MS11 and MS12 loci featured the highest deviation from Hardy-Weinberg equilibrium (HWE), with significant positive F_{IS} values in five and four populations, respectively. Locus MS6 and MS10 had significant negative F_{IS} values in four and one population, respectively. Overall, there was a heterozygote deficiency in all populations (except for SIA, with negative but non-significant F_{IS} value), irrespective of the nature of the substrate (Table 1).

Within locality spatial genetic structure

The kinship coefficient (Fig. 2A) and Rousset's $\hat{a}$ distance (Fig. 2B) for the different distance classes in the autocorrelation analyses showed a lack of pattern with

increasing distance in all four populations, with values remaining mostly within the confidence interval of no association (shaded areas in the figures). Only in a few cases the actual values of the coefficients fell slightly outside this no-association interval, which we attributed to random outcomes rather than to any true biological structure. This lack of intra-population structure allowed us to use the sampled specimens in inter-population comparisons (see below).

The differences between natural and artificial substrate were not significant neither for the overall kinship coefficient (Mann-Whitney U -test, $T = 15.5$; $P = 0.486$) nor for Rousset's $\hat{a}$ index ($T = 19$; $P = 0.886$).

Between population structure related to substrate type and geographical distance

Table 2 shows F_{ST} and D_{est} results. Both estimators' values were low in all cases. No pairwise F_{ST} comparison was significant after correction for multiple comparisons. Likewise, no pairwise D_{est} value was significantly different from 0, and neither was the global D_{est} value (0.0036). In addition, we computed the Pearson correlation between F_{ST} and D_{est} values to ensure that both estimators yielded comparable information, and the correlation was significant and very high ($r = 0.983$, $P < 0.001$). The Mantel test correlating the genetic and the geographic distances was not significant (using D_{est} matrix: $P = 0.977$; F_{ST} matrix: $P = 0.394$) so there was no indication of isolation by distance among these populations.

No significant variation attributable to differences between substrates (natural vs. artificial) was found in the AMOVA analysis (percent variation explained: 0.01%, $P = 0.425$, Table 3), neither among populations within substrate. The highest percentage of variation was explained by differences within individuals (89.64%, $P < 0.001$) and within populations (10.44%, $P < 0.001$).

Admixture analysis

The most likely number of genetically differentiated clusters was two as assessed by the programme STRUCTURE (IncK presented the highest value for $K = 2$, while the likelihood of the model presented the lowest standard deviation among the 20 replicates and a marked change in slope, which started to reach a plateau Fig. S1). In order to test whether the two clusters defined were influenced differentially by some loci, we ran STRUCTURE analyses with five loci, leaving out one locus at a time. This procedure showed that loci MS11 and MS12 were determinant of the genetic structure found (results not shown).

The two clusters were found in all eight populations studied (Fig. 3). The program STRUCTURE also gave us the probability of each individual to belong to each of the two genetically differentiated clusters. Interestingly, most individuals (85.1%) could be assigned to one of the two clusters with an assignment probability higher than 80%. A total of 151 individuals (henceforth group A) were assigned to one of the clusters and 106 individuals (henceforth group B) to the other cluster. Only 45 individuals (14.9%) had probabilities of assignment lower than 80% and were joined in a third group of individuals with mixed genotypes henceforth referred to as group AB. The procedure *find.clusters* of the *adeigenet* package with $K=2$ classified the individuals into two groups fully coherent with those of STRUCTURE: 83% of individuals of group A belonged to one of these groups, all individuals of group B belonged to the other and, as expected, the admixed individuals (AB) were classed almost equally (51:49%) into each group.

Table 4 summarizes the genetic information of the three groups established. Group AB presented the highest total allelic richness (33.86). Only group A showed a

significant (albeit low) F_{IS} value (0.097), which was due to the MS12 locus (F_{IS} = 0.796), and the overall F_{IS} values of the other two groups were close to zero and non-significant. 18 out of 45 alleles were exclusive of group A or B (irrespective of whether they were present in AB or not), but most of them were low frequency alleles (<1%) with the exception of loci MS11 and MS12, where four alleles with frequencies 4-32% were exclusive. Two alleles at low frequency (<4%) were found only in AB individuals.

The first axis of the FCA accounted for most of the variance (82.9%) and the second axis explained 17.0%. Axis 1 separated individuals assigned to the group A in the positive values, individuals assigned to the group B in the negative values, and individuals assigned to the group AB in between, the latter with considerable spread over the second axis (Fig. 4). Examining the relative contribution of alleles to the different axes, it is apparent that alleles from loci MS11 and MS12, as found in the STRUCTURE analyses, were the main contributors to the inertia explained by the first axis, followed by MS6. Likewise, MS12, MS10 and MS11 were the main contributors to axis 2.

We performed a discriminant analysis of principal components (DAPC) using the groups A, B and AB, together with individuals of two populations of Australia: Bunbury and Manly, representative of the two areas of origin of the worldwide introduced populations of *Microcosmus squamiger* (Rius et al. 2012) and genotyped for the same loci used in the present study. Along the first axis of the DAPC (explaining 57% of the total variance) group A appears close to Bunbury and group B to Manly, while admixed individuals (group AB) appeared in between (Fig. 5). The second axis (26% of variance) separates the native from the introduced populations. Additionally, we ran DAPC with only the Bunbury and Manly populations as groups and used the resulting discriminant function to assign individuals of groups A, B, and AB to one of

the two Australian populations using the function *predict.dapc*. A total of 78.14% of individuals of group A were assigned to Bunbury, 77.35% of group B were assigned to Manly, and individuals of group AB were evenly assigned to Bunbury (55.55%) or Manly (44.45%).

To ensure that the lack of inter-population genetic structure found in the previous analyses was not an artefact of the presence of individuals of different genetic pools mixed in all populations, we performed an AMOVA, grouping the individuals as per genetic group (A, B or AB) and examining the differentiation between populations within groups. A significant amount of the total variation (10.32%, $P < 0.001$) was explained by differences among the three genetic groups. Reassuringly, after accounting for this group structure, the residual variance between populations was low (0.11%) and not significant ($P = 0.369$) (Table 3), with most of the variability found within individuals (85.99%, $P < 0.001$). Likewise, we repeated the autocorrelation analyses separately for groups A and B within each population and did not find any significant fine-scale genetic structure (results not shown).

DISCUSSION

The present study revealed no spatial genetic structure at any of the scales analysed in populations of *Microcosmus squamiger* established on natural and artificial substrate, either within or between populations. No fine-scale (tens of metres) pattern was apparent and no significant genetic differentiation was found among pairs of populations (tens of km). Likewise, in none of the studied parameters did the populations on natural and artificial substrate differ.

Our expectation of lower genetic diversity and higher genetic relatedness in populations of *M. squamiger* on artificial substrates was not borne out, as allelic

richness, gene diversity and relatedness (kinship coefficient) were not different when the four populations on artificial substrate were compared with the four on natural rocky reefs. Likewise, grouping localities according to substrate type did not explain any significant amount of genetic variability in our samples (AMOVA), indicating that they likely represent a collective source of propagules to ensure the post-border dispersal success of this species.

Two differentiated genetic pools were identified with STRUCTURE. These results were coherent with the FCA ordination and with groups found using a *K*-means clustering algorithm based on a different approach (i.e, maximizing inter-group differentiation, Jombart et al. 2010). The separation of the two groups was due mostly to two of the loci studied (MS11 and MS12). This may reflect differences in allele frequencies between the two ancestral sources due to drift, although we cannot rule out that this outcome is due to differential selection on some genes linked to those two loci.

An unexpected pattern that appeared from the admixture analyses is that most individuals could be unambiguously assigned to one or another genetic group encountered, with only 15% of individuals showing ambiguous membership probabilities attributable to admixture between the two genetic pools. The general low frequency of private alleles of one group or the other may explain why admixed individuals do not have a particularly higher degree of heterozygotes than either of the groups. Moreover, even if not admixed at the individual level, the two genetic pools detected appeared mixed at the population level, and lack of spatial structure was also obtained when each identified genetic group was analysed separately. Remarkably, the groups seemed to correspond to an ancient split of the genetic makeup of the species, as they could be related to the two Australian populations of *M. squamiger* representative

of the two native areas that were reported as the source of the worldwide introduction of this species (Rius et al. 2012).

Many studies have addressed the potential relationship between spatial genetic structure at several scales and dispersal capabilities of larvae in benthic invertebrates (reviewed in Grosberg and Cunningham 2001; Hellberg et al. 2002; Shanks et al. 2003; Palumbi 2004). For example, fine-scale (within population) genetic structure can appear in species that are brooders and have extremely short-lived larvae such as sponges or cnidarians (e.g., Calderón et al. 2007; Blanquer et al. 2009; Ledoux et al. 2010; Mokhtar-Jamaï et al. 2011). All larval types of ascidians are short-lived and lecithotrophic, and solitary ascidians are commonly oviparous (Svane and Young 1989). In *M. squamiger*, time from fertilization to hatching is 11-12 h in the laboratory at 20°C for most embryos (Rius et al. 2010), although lower temperatures increase this time considerably (M.R. unpublished data). Once hatched, most settlement occurs during the first eight hours in laboratory conditions (Rius et al. 2010). Nonetheless, this potential lifespan may not correlate with natural dispersal if retention mechanisms occur, and several such mechanisms have been described for solitary ascidians including, among others, negative buoyancy, stickiness of eggs, and retention of eggs and larvae in mucus strings or bio-foam (Svane and Havenhand 1993; Petersen and Svane 1995; Marshall 2002; Castilla et al. 2007). In other solitary ascidians extremely localized dispersal has been reported (e.g., *Ciona intestinalis* Petersen and Svane 1995; Howes et al. 2007) and autocorrelation analyses have also substantiated the existence of genetic structure at a scale of a few meters (*Styela plicata*, David et al. 2010; *Styela clava*, Dupont et al. 2009). In *M. squamiger* no evidence of fine-scale structure exists in the range of distances analysed in the present study (up to ca. 100 meters), and the pattern is the

same in the populations studied on natural and artificial substrates. Therefore, no local retention mechanisms of embryos or larvae seem to be acting.

We have found a lack of genetic differentiation among our populations, indicating that enough larvae travel between our study localities separated by tens of kilometres so as to prevent the drifting apart of allele frequencies. This may reflect the natural active dispersal capabilities of this species but, even if we have chosen only open-shore sites, marinas are present all along the shore studied, and *M. squamiger* is present on them. Therefore, recreational boating is likely to have contributed to the movement of the species, which can then spill out of the confined environment of marinas and colonize nearby open-shore substrates. We cannot assess at present the relative importance of natural and man-mediated dispersal at the scale studied. Both have probably an important role in generating a pattern of high gene flow over natural and artificial substrates in the area, which are representative of the spacing of available hard substrate that is common in the highly urbanized western Mediterranean Sea (Airolidi and Beck 2007; Airolidi and Bulleri 2011).

The genetic exchange between populations is likely to be high and involve enough number of individuals as to avoid secondary bottlenecks in post-border dispersal. In our case, it is remarkable that in such a short stretch of coastline, 45 alleles were found in the six microsatellite loci, representing 71.4% of the total allelic richness found worldwide (63 alleles, Rius et al. 2012). This is in accordance with the source of the worldwide introduction of *M. squamiger* being traced to a single admixture event involving two genetically differentiated ancestral regions - the western and eastern coasts of Australia (Rius et al. 2012). However, in that work true admixture (interbreeding of the genetic pools) could not be distinguished from coexistence of the two unmixed groups in introduced populations. Our findings indicate that the

introduced populations studied have individuals attributable to one or the other putative ancestral genetic pools and that interbreeding among them is scarce, since only few individuals are actually admixed. The maintenance of this genetic structure is surprising, especially since historical records and genetic analyses show that the Mediterranean Sea was probably the first area of introduction of the species (Monniot 1981; Rius et al. 2012) more than 50 years ago, an ample time frame for the genetic pools to interbreed. Given the presence of two genetic clusters in each population, the overall deficit of heterozygotes detected and the significant F_{IS} values found in the four studied populations could be explained by a Wahlund effect. The alternative explanation for positive F_{IS} values, an artefact due to null alleles, seems unlikely since only MS10 had an appreciable percentage of failed amplifications (3.6 %) and this locus had more negative than positive F_{IS} values. Only one population (SIA) presented a non-significant negative F_{IS} value. However, it is also the population with the smallest sample size, which can influence this outcome.

We cannot tell at present which are the causes that prevent the interbreeding of the two genetic pools. Among other explanations, gamete recognition mechanisms or temporal separation of the breeding periods could be acting at the pre-zygotic stage. Previous studies of the life cycle of *M. squamiger* within our study area have shown a long reproductive period, from late spring to early autumn (Rius et al. 2009), which can accommodate differential reproductive timings. Reproductive post-zygotic isolation mechanisms reducing fitness of the admixed individuals are also a possibility. An alternative explanation would be that one of the groups (A or B) has recently arrived into the zone as a secondary introduction and little admixture has been realized since. Although this possibility cannot be discounted without a proper temporal sampling, it seems unlikely given that the Mediterranean was the first area colonized by a pool that

548 already comprised a mixture of the two native sources (Rius et al. 2012). Furthermore,
549 that study (whose sampling was done in 2006) included one population (Cubelles) in
550 common with the present work. A reanalysis of that population (done with
551 STRUCTURE with $K=2$) showed that the two genetic groups (A and B) here found
552 were already present in 2006 with only a few (16%) admixed individuals (AB). Another
553 potential explanation for the presence of two genetic pools, cryptic speciation, can be
554 discarded using the sequences of COI for the Cubelles population sampled in 2006
555 available in Rius et al. (2008a). The sequences of the majority (70%) of individuals
556 were included in the most abundant haplogroup in the introduced range (Rius et al.
557 2008a), irrespective of whether the individuals belonged to groups A, B, or AB.

558 Whatever the proximate cause, the maintenance over time of differentiated
559 genetic pools can have a strong effect on the introduction process itself. It has been
560 recognized that multiple introductions, if originating from different sources, contribute
561 to propagule pressure and, perhaps more importantly, to the generation of novel allelic
562 combinations that can foster the adaptive potential of invasive populations (Geller et al.
563 2010). In another introduced solitary ascidian, *Styela plicata*, two ancestral genetic
564 groups were also recognized but the populations had in general an excess of admixed
565 genotypes with respect to expectation (Pineda et al. 2011), so that an interbreeding of
566 the two genetic pools seemed to be favoured in this case. Although not demonstrated
567 empirically, admixture has been suggested to increase the ability of invasive species to
568 adapt and to thrive in new environments in terrestrial, freshwater, and marine systems
569 (e.g. McIvor et al. 2001; Kelly et al. 2006; Blum et al. 2007; Facon et al. 2008; Kolbe et
570 al. 2008; Lombaert et al. 2011). In *M. squamiger*, however, admixture seems to be
571 limited by unknown factors, without having compromised the colonizing potential of
572 the species. In any case, the existence of two distinct pools that may differ in

ecophysiological characteristics (e.g., reproductive timing) is undoubtedly to be kept in mind when studying the introduction process and the invasive potential of *M. squamiger*.

Taken together, our results indicate that the expansion of *M. squamiger* is highly likely across different substrate types available for colonization. Mitigation measures are thus difficult to undertake at the post-border stage in shores with enough patches of hard substrate. Efforts would be better invested in preventing the arrival of this species at the pre-border stage, a consideration that seems obvious for species with wide dispersal abilities and that, considering our results, applies also to organisms with short-lived larval stages.

ACKNOWLEDGMENTS

We thank the staff at the Scientific and Technical Services of the University of Barcelona for their help with the genotyping process. We are also grateful to three anonymous reviewers for constructive criticism that helped to improve the manuscript. This research was funded by projects CTM2010-22218 of the Spanish Government, BIOCON08-187 of the FBBVA, and PIE 200730I026 of the Spanish Research Council (CSIC). Funding came also from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° PIOF-GA-2009-254634 to M.R. V.O. was supported by a FPU PhD scholarship (AP2008 / 04209) from the Spanish Ministerio de Educación. The authors are part of the research groups 2009SGR-636, and 2009SGR-484 of the Generalitat de Catalunya (Spain).

REFERENCES

- Airolidi L, Beck MW (2007) Loss, status and trends for coastal marine habitats of Europe. *Oceanogr Mar Biol* 45: 345-405
- Airolidi L, Bulleri F (2011) Anthropogenic disturbance can determine the magnitude of opportunistic species responses on marine urban infrastructures. *PLoS ONE* 6: e22985. DOI: 10.1371/journal.pone.0022985
- Bacchiocchi F, Airolidi L (2003) Distribution and dynamics of epibiota on hard structures for coastal protection. *Estuar Coast Shelf S* 56: 1157-1166
- Barros R, Rocha R, Pie M (2009) Human-mediated global dispersion of *Styela plicata* (Tunicata, Ascidiacea). *Aquat Inv* 4: 45–57
- Bax N, Carlton JT, Mathews-Amos A, Haedrich RL, Howarth FG, Purcell JE, Rieser A, Gray A (2001) The control of biological invasions in the world's oceans. *Conserv Biol* 15: 1234–1246
- Belkhir K, Borsa P, Chikhi P, Raufaste N, Bonhomme F (2004) GENETIX 4.05.2, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II, Montpellier, France

622 Blackburn TM, Pysek P, Bacher S, Carlton JT, Duncan RP, Jarosik V, Wilson JRU,
 623 Richardson DM (2011) A proposed unified framework fo biological invasions. TREE
 624 26: 333-339
 625
 626 Blanquer A, Uriz M-J, Caujapé-Castells J (2009) Small-scale spatial genetic structure in
 627 *Scopalina lophyropoda*, an encrusting sponge with philopatric larval dispersal and
 628 frequent fission and fusion events. Mar Ecol-Prog Ser 380: 95-102
 629
 630 Blum JC, Chang AL, Liljesthröm M, Schenk ME, Steinberg MK, Ruiz GM (2007) The
 631 non-native solitary ascidian *Ciona intestinalis* (L.) depresses species richness. J Exp
 632 Mar Biol Ecol 342: 5-14
 633
 634 Branch GM, Steffani CN (2004) Can we predict the effects of alien species? A case-
 635 history of the invasion of South Africa by *Mytilus galloprovincialis* (Lamarck). J Exp
 636 Mar Biol Ecol 300: 189-215
 637
 638 Bulleri F, Aioldi L (2005) Artificial marine structures facilitate the spread of a non-
 639 indigenous green alga, *Codium fragile* ssp. *tomentosoides*, in the north Adriatic Sea. J
 640 Appl Ecol 42: 1063-1072
 641
 642 Bulleri F, Chapman MG (2010) The introduction of coastal infrastructure as a driver of
 643 change in marine environments. J Appl Ecol 47: 26–35
 644

645 Calderón I, Ortega N, Durán S, Pascual M, Turon X (2007) Finding the relevant scale:
 646 clonality and genetic structure in a marine invertebrate (*Crambe crambe*, Porifera). Mol
 647 Ecol 16: 1799–1810
 648
 649 Carlton JT, Geller JB (1993) Ecological roulette: the global transport of nonindigenous
 650 marine organisms. Science 261: 78-82
 651
 652 Castilla JC, Manríquez PH, Delgado AP, Gargallo L, Leiva A, Radic D (2007) Bio-
 653 foam enhances larval retention in a free-spawning marine tunicate. P Natl Acad Sci
 654 USA 104: 18120-18122
 655
 656 David GK, Marshall DJ, Riginos C (2010) Latitudinal variability in spatial genetic
 657 structure in the invasive ascidian, *Styela plicata*. Mar Biol 157: 1955-1965
 658
 659 Dijkstra J, Harris LG, Westeman E (2007) Distribution and long-term temporal patterns
 660 of four invasive colonial ascidians in the Gulf of Maine. J Exp Mar Biol Ecol 342: 61–
 661 68
 662
 663 Dumont CP, Gaymer CF, Thiel M (2011) Predation contributes to invasion resistance of
 664 benthic communities against the non-indigenous tunicate *Ciona intestinalis*. Biol Inv
 665 13: 2023-2034
 666
 667 Dupont L, Viard F, Dowell MJ, Wood C, Bishop JDD (2009) Fine- and regional-scale
 668 genetic structure of the exotic ascidian *Styela clava* (Tunicata) in southwest England, 50
 669 years after its introduction. Mol Ecol 18: 442-453

670

671 Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software
672 package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:
673 47-50

674

675 Evanno C, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals
676 using the software STRUCTURE: a simulation study. *Mol Ecol* 14: 2611-2620

677

678 Fauvelot C, Bertozzi F, Costantini F, Airoidi L, Abbiati M (2009) Lower genetic
679 diversity in the limpet *Patella caerulea* on urban coastal structures compared to natural
680 rocky habitats. *Mar Biol* 156: 2313-2323

681

682 Fenster CB, Vekemans X, Hardy OJ (2003) Quantifying gene flow from spatial genetic
683 structure data in a metapopulation of *Chamaecrista fasciculata* (Leguminosae).
684 *Evolution* 57: 995–1007

685

686 Facon B, Pointier JP, Jarne P, Sarda V, David P (2008) High genetic variance in life-
687 history strategies within invasive populations by way of multiple introductions. *Curr*
688 *Biol* 18: 363-367

689

690 Forrest BM, Gardner JPA, Taylor MD (2009) Internal borders for managing invasive
691 marine species. *J Appl Ecol* 46: 46-54

692

693 Geller JB, Darling JA, Carlton JT (2010) Genetic perspectives on marine biological
694 invasions. *Annu Rev Mar Sci* 2: 367–393

695

696 Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P (2010) Calculations of
697 population differentiation based on G_{ST} and D : forget G_{ST} but not all of statistics!
698 Mol Ecol 19: 3845-3852

699

700 Glasby TM, Connell SD, Holloway MG, Hewitt CL (2007) Nonindigenous biota on
701 artificial structures: could habitat creation facilitate biological invasions? Mar Biol 151:
702 887-895

703

704 Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillis RW, Benz WK,
705 Robertson HM, Engels WR (1993). Type I repressors of P element mobility. Genetics
706 135: 81-95

707

708 Goldstien SJ, Schiel DR, Gemmell NJ (2010) Regional connectivity and coastal
709 expansion: differentiating pre-border and post-border vectors for the invasive tunicate
710 *Styela clava*. Mol Ecol 19: 874–885

711

712 Goldstien SJ, Dupont L, Viard F, Hallas PJ, Nishikawa T, et al. (2011) Global
713 Phylogeography of the Widely Introduced North West Pacific Ascidian *Styela clava*.
714 PLoS ONE 6: e16755. doi:10.1371/journal.pone.0016755

715

716 Goudet J (2002) FSTAT version 2.9.3.2. Available from Jerome.goudet @ie.zea.unil.ch,
717 via email. Institute of Ecology, UNIL, CH-1015, Lausanne, Switzerland

718

719 Grosberg R, Cunningham CW (2001) Genetic structure in the sea. From populations to
720 communities. In: Bertness MD, Gaines SD, Hay ME, eds. Marine Community Ecology
721 Sinauer Associates, Inc., Sunderland, Massachusetts. pp 61–84
722

723 Hardy OJ, Vekemans X (1999) Isolation by distance in a continuous population:
724 reconciliation between spatial autocorrelation analysis and population genetics models.
725 Heredity 83: 145–154
726

727 Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse
728 spatial genetic structure at the individual or population levels. Mol Ecol Notes 2: 618-
729 620
730

731 Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59: 1633-
732 1638
733

734 Hellberg ME, Burton RS, Neigel JE, Palumbi SR (2002) Genetic assessment of
735 connectivity among marine populations. B Mar Sci 70: 273-290
736

737 Heywood JS (1991) Spatial analysis of genetic variation in plant populations. Annu Rev
738 Ecol Syst 22: 335-355
739

740 Howes S, Herbinger CM, Darnell P, Vercaemer B (2007) Spatial and temporal patterns
741 of recruitment of the tunicate *Ciona intestinalis* on a mussel farm in Nova Scotia,
742 Canada. J Exp Mar Biol Ecol 342: 85-92
743

744 Hubisz M, Falush D, Stephens M, Pritchard J (2009) Inferring weak population
 745 structure with the assistance of sample group information. *Mol Ecol Res* 9: 1322-1332
 746
 747 Hulme PE (2006) Beyond control: wider implications for the management of biological
 748 invasions. *J Appl Ecol* 43: 835-847
 749
 750 Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation
 751 program for dealing with label switching and multimodality in analysis of population
 752 structure. *Bioinformatics* 23: 1801-1806
 753
 754 Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic
 755 markers. *Bioinformatics* 24: 1403–1405
 756
 757 Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal
 758 components: a new method for the analysis of genetically structured populations. *BMC*
 759 *Genetics* 11: 94. doi:10.1186/1471-2156-11-94
 760
 761 Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Mol Ecol* 17: 4015–
 762 4026
 763
 764 Kelly DW, Muirhead JR, Heath DD, MacIsaac HJ (2006) Contrasting patterns in
 765 genetic diversity following multiple invasions of fresh and brackish waters. *Mol Ecol*
 766 15: 3641-3653
 767

768 Kolbe JJ, Larson A, Losos JB, de Queiroz K (2008) Admixture determines genetic
 769 diversity and population differentiation in the biological invasion of a lizard species.
 770 Biol Letters 4: 434-437
 771
 772 Kott P (1985) The Australian Ascidiacea Pt 1, Phlebobranchia and Stolidobranchia.
 773 Mem Queens Mus 23: 1-440
 774
 775 Lacoursière-Roussel A, Bock DG, Cristescu ME, Guichard F, Girard P, Legendre P,
 776 McKindsey CW (2012) Disentangling invasion processes in a dynamic shipping-
 777 boating network. Mol Ecol 21: 4227-4241
 778
 779 Lambert G (2002) Nonindigenous Ascidiarians in Tropical Waters. Pac Sci 56: 291-298
 780
 781 Lambert G (2007) Invasive sea squirts: A growing global problem. J Exp Mar Biol Ecol
 782 342: 3-4
 783
 784 Lambert CC, Lambert G (1998) Non-indigenous ascidians in southern California
 785 harbors and marinas. Mar Biol 130: 675-688
 786
 787 Lambert CC, Lambert G (2003) Persistence and differential distribution of nonindigenous
 788 ascidians in harbors of the Southern California Bight. Mar Ecol-Prog Ser 259: 145-161
 789
 790 Ledoux JB, Garrabou J, Bianchimani O, Drap P, Féral JP, Aurelle D (2010) Fine-scale
 791 genetic structure and inferences on population biology in the threatened Mediterranean
 792 red coral, *Corallium rubrum*. Mol Ecol 19: 4202-4216

793

794 Lejeusne C, Bock DG, Therriault TW, MacIsaac HJ, Cristescu ME (2011) Comparative
795 phylogeography of two colonial ascidians reveals contrasting invasions histories in
796 North America. *Biol Inv* 13: 635-650

797

798 Locke A (2009) A screening procedure for potential tunicate invaders of Atlantic
799 Canada. *Aquat Inv* 4: 71-79

800

801 Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical
802 understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* 82: 1420–1425

803

804 Lombaert E, Guillemaud T, Thomas CE, Lawson Handley LJ, Li J, Wang S, Pang H,
805 Goryacheva I, Zakharov IA, Jousset E, Poland RL, Migeon A, van Lenteren J, De
806 Clercq P, Berkvens N, Jones W, Estoup A (2011) Inferring the origin of populations
807 introduced from a genetically structured native range by approximate Bayesian
808 computation: case study of the invasive ladybird *Harmonia axyridis*. *Mol Ecol* 20:
809 4654-4670

810

811 López-Legentil S, Turon X, Planes S (2006) Genetic structure of the star sea squirt,
812 *Botryllus schlosseri*, introduced in southern European harbours. *Mol Ecol* 15: 3957–
813 3967

814

815 Marshall DJ (2002) In situ measures of spawning synchrony and fertilization success in
816 an intertidal, free-spawning invertebrate. *Mar Ecol-Prog Ser* 236: 113-119

817

818 McIvor L, Maggs CA, Provan J, Stanhope J (2001) *rbcL* sequences reveal multiple
819 cryptic introductions of the Japanese red alga *Polysiphonia harveyi*. *Mol Ecol* 10: 911-
820 919
821

822 Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related
823 measures. *Mol Ecol Res* 11: 5–18
824

825 Meusnier I, Valero M, Destombe C, Godé C, Desmarais E, Bonhomme F, Stam WT,
826 Olsen JL (2002) Polymerase chain reaction-single strand conformation polymorphism
827 analyses of nuclear and chloroplast DNA provide evidence for recombination, multiple
828 introductions and nascent speciation in the *Caulerpa taxifolia* complex. *Mol Ecol* 11:
829 2317-2325
830

831 Michaelsen VW (1927) Einige neue westaustralische Ptychobranchiate Ascidien. *Zool*
832 *Anz* 71: 193-203
833

834 Mokhtar-Jamaï K, Pascual M, Ledoux JB, Coma R, Féral JP, Garrabou J, Aurelle D
835 (2011) From global to local genetic structuring in the red gorgonian *Paramuricea*
836 *clavata*: the interplay between oceanographic conditions and limited larval dispersal.
837 *Mol Ecol* 20: 3291-3305
838

839 Monniot C (1981) Apparition de l'ascidie *Microcosmus exasperatus* dans les ports
840 Méditerranéens. *Téthys* 10: 59-62
841

842 Narum SR (2006) Beyond Bonferroni: Less conservative analyses for conservation
 843 genetics. *Conserv Genet* 7: 783–787
 844

845 Osman RW, Whitlatch RB (2007) Variation in the ability of *Didemnum* sp. to invade
 846 established communities. *J Exp Mar Biol Ecol* 342: 40-53
 847

848 Palumbi SR (2004) Marine reserves and ocean neighborhoods: the spatiel scale of
 849 marine populations and their management. *Ann Rev Env Res* 29: 31-68
 850

851 Park SDE (2001) Trypanotolerance in West African Cattle and the Population Genetic
 852 Effects of Selection [Ph.D. thesis] University of Dublin
 853

854 Pérez-Portela R, Turon X, Bishop JDD (2012) Bottlenecks and loss of genetic diversity:
 855 spatio-temporal patterns of genetic structure in an ascidian recently introduced in
 856 Europe. *Mar Ecol-Prog Ser* 451: 93-105
 857

858 Petersen JK, Svane I (1995) Larval dispersal in the ascidian *Ciona intestinallis* (L.).
 859 Evidence for a closed population. *J Exp Mar Biol Ecol* 186: 89-102
 860

861 Pineda MC, López-Legentil S, Turon X (2011) The whereabouts of an ancient
 862 wanderer: global phylogeography of the solitary ascidian *Styela plicata*. *PLoS ONE* 6:
 863 e25495. doi:10.1371/journal.pone.0025495
 864

865 Pinn EH, Mitchell K, Corkill J (2005) The assemblages of groynes in relation to
 866 substratum age, aspect and microhabitat. *Estuar Coast Shelf S* 62: 271-282

867

868 Piry S, Alapetite A, Cournet JM, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2:
869 A Software for Genetic Assignment and First-Generation Migrant Detection. J Hered
870 95: 536-539

871

872 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using
873 multilocus genotype data. Genetics 155: 945-959

874

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

877

878 Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding
879 coefficients. Genet Res 67: 175–185

880

881 Rius M, Pascual M, Turon X (2008a) Phylogeography of the widespread marine invader
882 *Microcosmus squamiger* (Ascidacea) reveals high genetic diversity of introduced
883 populations and non-independent colonizations. Divers Dist 14: 818-828

884

885 Rius M, Turon X, Pascual M (2008b) Isolation of polymorphic microsatellite loci for
886 the marine invader *Microcosmus squamiger* (Ascidacea). Mol Ecol Res 8: 1405-1407

887

888 Rius M, Pineda MC, Turon X. 2009. Population dynamics and life cycle of the
889 introduced ascidian *Microcosmus squamiger* in the Mediterranean Sea. Biol Inv 11:
890 2181-2194

891

892 Rius M, Turon X, Muniz-Dias G, Marshall DJ (2010) Propagule size effects across
893 multiple life-history stages in a marine invertebrate. *Funct Ecol* 24: 685-693
894

895 Rius M, Turon X, Ordóñez V, Pascual M. (2012) Tracking invasion histories in the sea:
896 facing complex scenarios using multilocus data. *PLoS ONE* 7: e35815.
897 doi:10.1371/journal.pone.0035815
898

899 Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic
900 invasions. *TREE* 22: 454-464
901

902 Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population
903 structure. *Mol Ecol Notes* 4: 137-138
904

905 Rousset F (2000) Genetic differentiation between individuals. *J Evolution Biol* 13: 58–
906 62
907

908 Ruiz GM, Carlton JT, Grosholz ED, Hines AH (1997) Global invasions of marine and
909 estuarine habitats by non-indigenous species: mechanisms, extent, and consequences.
910 *Amer Zool* 37: 621-632
911

912 Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S,
913 Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN,
914 Weller SG (2001) The population biology of invasive species. *Annu Rev Ecol Syst* 32:
915 305-332
916

917 Shanks AL, Grantham BA, Carr MH (2003) Propagule dispersal distance and the size
 918 and spacing of marine reserves. *Ecol Appl* 13: S159-S169
 919
 920 Simkanin C, Davidson IC, Dower JF, Jamieson G, Therriault TW (2012) Anthropogenic
 921 structures and the infiltration of natural benthos by invasive ascidians. *Mar Ecol* 33:
 922 499-511
 923
 924 Stachowicz JJ, Tilman D (2005) Species invasions and the relationships between
 925 species diversity, community saturation, and ecosystem functioning. In: Sax DF,
 926 Stachowicz JJ, Gaines SD (eds) *Species Invasions: Insights into Ecology, Evolution and*
 927 *Biogeography*, Sinauer Associates, Inc., Sunderland, pp 41–64
 928
 929 Svane I, Havenhand JN (1993) Spawning and dispersal in *Ciona intestinallis* (L.). *Mar*
 930 *Ecol* 14: 53-66
 931
 932 Svane I, Young CM (1989) The ecology and behaviour of ascidian larvae. *Oceanogr*
 933 *Mar Biol: Annu Rev* 27: 45-90
 934
 935 Turon X, Rius M, Nishikawa T (2007) Spread of *Microcosmus squamiger* (Ascidacea:
 936 Pyuridae) in the Mediterranean Sea and adjacent waters. *J Exp Mar Biol Ecol* 342: 185-
 937 188
 938
 939 Tyrrell MC, Byers JE (2007) Do artificial substrates favor nonindigenous fouling
 940 species over native species? *J Exp Mar Biol Ecol* 342: 54-60
 941

942 Vaselli S, Bulleri F, Benedetti-Cecchi L (2008) Hard coastal-defence structures as
 943 habitats for native and exotic rocky-bottom species. *Mar Environ Res* 66: 395-403
 944

945 Wasson K, Zabin CJ, Bedinger, Diaz MC, Pearse JS (2001) Biological invasions of
 946 estuaries without international shipping: the importance of intraregional transport. *Biol*
 947 *Conserv* 102: 143-153
 948

949 Weir BS, Cockerham CC (1984) Estimating F-Statistics for the analysis of population
 950 structure. *Evolution* 38: 1358-1370
 951

952 Whitlock MC (2011) G'_{ST} and D do not replace F_{ST} . *Mol Ecol* 20: 1083-1091
 953

954 Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, Richardson DM (2009) Something in
 955 the way you move: dispersal pathways affect invasion success. *TREE* 24: 136-144
 956

957 Wonham MJ, Walton WC, Ruiz GM, Frese AM, Galil BS (2001) Going to the source:
 958 role of the invasion pathway in determining potential invaders. *Mar Ecol-Prog Ser* 215:
 959 1-12
 960

961 Wotton DM, Hewitt CL (2004) Marine biosecurity post-border management:
 962 developing incursion response systems for New Zealand. *New Zeal J Mar Fresh Res* 38:
 963 553-559
 964

965 Zhan A, MacIsaac HJ, Cristescu ME (2010) Invasion genetics of the *Ciona intestinalis*
966 species complex: from regional endemism to global homogeneity. Mol Ecol 19: 4678-
967 4694
968
969 Zhan A, Darling JA, Bock DG, Lacoursière-Roussel A, MacIsaac HJ, Cristescu ME
970 (2012) Complex genetic patterns in closely related colonizing invasive species. Ecol
971 Evol 2: 1331-1346
972

Fig. 1 Sampled localities between the ports of Barcelona and Tarragona (Spain, northwestern Mediterranean). In **bold** the localities sampled on natural substrate and in *italics* the localities sampled on artificial substrate. Population codes as in Table 1.

A)

B)

Fig. 2 Graphic representation of the A) kinship coefficient and B) Rousset's $\hat{a}$ index at each distance class in populations on natural substrate (Sitges and El Roc de Sant Gaietà) and artificial substrate (Cubelles and Segur de Calafell). The first distance class (0.2) corresponds to individuals collected in the different sampling points within a 20 cm radius. The shaded areas represent the 95% confidence interval of non-association between genetic and spatial structure obtained by randomly permuting individuals across positions of the transects. Codes of localities as in Table 1.

Fig. 3 Population structure in the eight populations studied with the most likely number of clusters ($K = 2$) inferred with the program STRUCTURE. Population codes as in Table 1.

1045

1046

1047

1048

1049

1050

1051

1052

1053

1054

1055

1056

Fig. 4 Two-dimension plot of the Factorial Correspondence Analysis (FCA) of the 302 individuals. The three symbols indicate in which group (A, B, AB) each individual belongs.

Fig. 5 Discriminant analysis of principal components (DAPC) using as groups the two Australian native populations: Bunbury (BU) and Manly (MA) and the three groups of individuals (A, B, AB) defined after the STRUCTURE results within the introduced populations studied herein.

Table 1. Genetic variability of natural and artificial substrate populations. Number of amplified individuals (N); number of alleles (NA), private alleles (if any) are indicated inside parenthesis; allelic richness per locus and population (AR), based on a minimum amplified sample size (over all loci) of 15 diploid individuals; observed (Ho) and expected (He) heterozygosities; and inbreeding coefficient (F_{IS}). Significant F_{IS} values are in **bold**. Means over loci (or global value for F_{IS}) are also indicated.

		Locus						Mean
		MS 6	MS 7	MS 10	MS 11	MS 12	MS 13	
Populations on Natural Substrate								
Garraf (GAN)								
N	18	18	16	18	18	18		17.667
NA	2	2	5	6	5	4		4.000
AR	2.000	2.000	4.938	5.833	4.807	3.833		3.902
Ho	0.278	0.389	0.625	0.667	0.333	0.556		0.475
He	0.386	0.475	0.613	0.795	0.502	0.663		0.572
<i>F_{IS}</i>	0.286	0.185	-0.020	0.166	0.342	0.167		0.175
Sitges (SIN)								
N	59	59	59	59	59	59		59.000
NA	2	3 (1)	7	6	6	4		4.667
AR	1.997	2.254	4.853	5.776	5.174	3.694		3.958
Ho	0.271	0.576	0.712	0.441	0.254	0.627		0.480
He	0.261	0.458	0.621	0.792	0.666	0.646		0.574
<i>F_{IS}</i>	-0.040	-0.262	-0.148	0.446	0.621	0.030		0.165
El Roc de Sant Gaietà (SG)								
N	45	45	44	45	45	45		44.833
NA	3 (1)	2	7	7	7	5 (1)		5.167
AR	2.331	2.000	5.544	5.873	5.750	4.137		4.272
Ho	0.333	0.511	0.614	0.622	0.444	0.556		0.513
He	0.284	0.425	0.604	0.800	0.622	0.651		0.564
<i>F_{IS}</i>	-0.174	-0.206	-0.016	0.224	0.287	0.148		0.091
Torredembarra (TB)								
N	34	34	34	34	34	34		34.000
NA	3	2	7	6	8	4		5.000
AR	2.441	2.000	5.425	5.883	5.892	3.685		4.221
Ho	0.353	0.500	0.588	0.676	0.441	0.559		0.520
He	0.355	0.409	0.547	0.808	0.543	0.595		0.543
<i>F_{IS}</i>	0.006	-0.228	-0.076	0.165	0.191	0.062		0.043
Populations on Artificial Substrate								
Garraf (GAA)								
N	27	27	24	27	27	27		26.500
NA	2	2	8	7	5	4		4.667
AR	1.998	2.000	6.124	6.869	4.803	3.913		4.284
Ho	0.259	0.296	0.708	0.630	0.296	0.593		0.464
He	0.230	0.425	0.621	0.840	0.584	0.635		0.556
<i>F_{IS}</i>	-0.130	0.307	-0.143	0.254	0.498	0.068		0.168
Sitges (SIA)								
N	16	16	15	16	16	16		15.833
NA	2	2	3	6	3	4		3.333
AR	2.000	2.000	3.000	5.935	3.000	3.938		3.312
Ho	0.438	0.500	0.267	0.813	0.750	0.750		0.586
He	0.353	0.484	0.246	0.772	0.575	0.688		0.520

<i>F_{IS}</i>	-0.250	-0.034	-0.087	-0.054	-0.319	-0.094	-0.133
Cubelles (CU)							
N	56	57	57	57	56	57	56.667
NA	2	2	11	7 (1)	10 (1)	5 (1)	6.167
AR	1.999	2.000	6.752	6.714	6.616	4.150	4.705
Ho	0.357	0.474	0.579	0.649	0.446	0.596	0.517
He	0.296	0.429	0.580	0.834	0.700	0.613	0.575
<i>F_{IS}</i>	-0.209	-0.104	0.001	0.223	0.364	0.028	0.102
Segur de Calafell (SC)							
N	46	46	42	46	45	46	45.167
NA	4 (1)	2	8	6	8	4	5.333
AR	2.650	2.000	6.415	5.896	6.075	3.698	4.456
Ho	0.304	0.326	0.548	0.587	0.378	0.717	0.477
He	0.297	0.379	0.618	0.818	0.527	0.656	0.549
<i>F_{IS}</i>	-0.024	0.141	0.116	0.284	0.285	-0.094	0.133

1076
1077
1078
1079
1080

Table 2. Genetic differentiation of the studied populations of *Microcosmus squamiger*. F_{ST} values are shown above the diagonal and D_{est} values below. Shaded cells indicate pairwise comparisons between natural and artificial substrate. In addition, mean values of each estimator ($\pm$ SE) are provided for comparison between populations on natural substrate (Nat-Nat), between natural and artificial substrate (Nat-Art) and between artificial substrate (Art-Art). Population codes as in Table 1.

	GAN	SIN	SG	TB	GAA	SIA	CU	SC
GAN		-0.0079	-0.0113	-0.0134	-0.0058	0.0092	-0.0009	-0.0074
SIN	-0.0100		-0.0072	0.0029	-0.0041	0.0202	0.0007	0.0025
SG	-0.0170	-0.0102		-0.0056	-0.0072	0.0141	-0.0038	-0.0059
TB	-0.0198	0.0071	-0.0057		-0.0018	0.0172	-0.0006	-0.0061
GAA	-0.0021	0.0013	-0.0052	-0.0006		0.0252	-0.0059	-0.0077
SIA	0.0081	0.0265	0.0150	0.0249	0.0363		0.0214	0.0192
CU	0.0023	0.0023	-0.0042	0.0028	-0.0079	0.0354		0.0026
SC	-0.0088	0.0055	-0.0066	-0.0074	-0.0114	0.0254	0.0041	

D_{est}	mean	$\pm$ SE
Nat-Nat	-0.0093	0.0039
Nat-Art	0.0034	0.0031
Art-Art	0.0137	0.0088

F_{ST}	mean	$\pm$ SE
Nat-Nat	-0.0071	0.0023
Nat-Art	0.0013	0.0026
Art-Art	0.0091	0.0059

Table 3. AMOVA considering two groups of populations as per substrate type (natural and artificial), and considering the three genetic groups herein defined from the STRUCTURE results (A, B, AB).

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	P-value
AMOVA by Type of Substrate					
<i>Between Substrates</i>	1	1.736	<0.001	0.010	0.425
<i>Among Populations within Substrates</i>	6	10.360	-0.001	-0.090	0.613
<i>Among individuals within Populations</i>	294	538.927	0.173	10.440	0.000
<i>Within individuals</i>	302	449	1.487	89.640	0.000
AMOVA by Genetic Groups					
<i>Between groups</i>	2	68.506	0.178	10.320	0.000
<i>Among populations within groups</i>	20	33.155	0.002	0.110	0.369
<i>Among individuals within populations</i>	279	449.362	0.062	3.580	0.019
<i>Within individuals</i>	302	449	1.487	85.990	0.000

Table 4. Genetic variability of the three groups (A, B, AB) defined following STRUCTURE results. Number of amplified individuals (N); number of alleles (NA); allelic richness per locus and population (AR) based on a minimum amplified sample size (over all loci) of 43 diploid individuals; observed (Ho) and expected (He) heterozygosities; and inbreeding coefficient (F_{IS}). Significant F_{IS} values are in **bold**. Means over loci (or global value for F_{IS}) are also indicated.

		Locus						Mean
		MS 6	MS 7	MS 10	MS 11	MS 12	MS 13	
A	N	150	151	144	151	150	151	149.500
	NA	4	3	10	6	3	6	5.333
	AR	2.573	2.285	7.79	5.955	2.871	4.564	4.339
	Ho	0.347	0.437	0.625	0.649	0.06	0.583	0.450
	He	0.332	0.437	0.598	0.697	0.294	0.633	0.499
	<i>F_{IS}</i>	-0.045	0.000	-0.046	0.069	0.796	0.079	0.097
	B	N	106	106	104	106	106	106
NA		3	2	11	7	7	4	5.667
AR		2.406	2	8.333	6.275	6.224	3.997	4.872
Ho		0.283	0.519	0.567	0.557	0.755	0.651	0.555
He		0.252	0.427	0.529	0.718	0.759	0.642	0.555
<i>F_{IS}</i>		-0.123	-0.216	-0.073	0.226	0.005	-0.014	-0.001
AB		N	45	45	43	45	44	45
	NA	3	2	10	7	8	4	5.667
	AR	2.956	2	10	6.997	7.954	3.956	5.644
	Ho	0.311	0.4	0.628	0.578	0.659	0.644	0.537
	He	0.269	0.396	0.674	0.741	0.563	0.632	0.546
	<i>F_{IS}</i>	-0.159	-0.011	0.069	0.223	-0.174	-0.019	0.017

Table S1. Populations studied, with code names and coordinates (Latitude and Longitude); type of substrate (S): natural (Nat) or artificial (Art); number of individuals sampled from each population (N).

Population	S	Code	Latitude	Longitude	N
Garraf	Nat	GAN	41°15'8.85"N	1°54'9.15"E	18
Sitges	Nat	SIN	41°14'11.16"N	1°49'45.64"E	59
El Roc de Sant Gaietà	Nat	SG	41°10'6.97"N	1°28'55.13"E	45
Torredembarra	Nat	TB	41° 8'46.19"N	1°25'9.09"E	34
Garraf	Art	GAA	41°15'8.19"N	1°54'8.37"E	27
Sitges	Art	SIA	41°14'8.99"N	1°49'39.39"E	16
Cubelles	Art	CU	41°11'37.93"N	1°39'12.08"E	57
Segur de Calafell	Art	SC	41°11'16.95"N	1°36'30.04"E	46

Table S2. Point distances and number of individual sampled in the transects laid out for the autocorrelation study in natural (Sitges and El Roc de Sant Gaietà) and artificial substrate populations (Cubelles and Segur de Calafell). In each transect point all individuals were collected within a 20 cm radius.

Distance	NATURAL		ARTIFICIAL	
	Sitges	El Roc de Sant Gaietà	Cubelles	Segur de Calafell
0m	9	6	4	6
2.5m	9	5	4	6
5m	9	6	9	5
10m	7	6	9	7
20m	9	7	9	6
40m	6	4	7	6
90m	10	11	15	10
Total	59	45	57	46

Fig. S1 Change in likelihood of the model ($\ln P(D) \pm \text{SD}$) (continuous line) and IncK (dashed line) statistic for increasing numbers of differentiated genetic clusters (K).