Mar Biodiv
DOI 10.1007/s12526-017-0672-9

SENCKENBERG

ORIGINAL PAPER

Novel microsatellite markers for *Pyura chilensis* reveal fine-scale genetic structure along the southern coast of Chile

E. C. Giles[1] · C. Petersen-Zúñiga[1] · S. Morales-González[1] · S. Quesada-Calderón[1] · Pablo Saenz-Agudelo[1]

Received: 25 August 2016 / Revised: 22 February 2017 / Accepted: 23 February 2017

Abstract Studying the geographic scale of gene flow and population structure in marine populations can be a powerful tool with which to infer patterns of larval dispersal averaged across generations. Here, we describe the development of ten novel polymorphic microsatellite markers for an important endemic ascidian, *Pyura chilensis*, of the southeastern Pacific, and we report the results from fine-scale genetic structure analysis of 151 *P. chilensis* individuals sampled from five sites constituting ~80 km of coastline in southern Chile. All microsatellite markers were highly polymorphic (number of alleles ranged from 12 to 36). Our results revealed significant deviations from Hardy–Weinberg equilibrium (HWE) for most loci, suggesting the presence of either null alleles or deviations from random mating within sampled sites. However, we found a significantly higher spatial autocorrelation and higher mean pairwise relatedness among individuals sampled from the same sites than would be expected if samples were randomly distributed across all sites; this suggests that spatial configuration and reproduction might not be random within sites. Our results indicate the presence of a weak but significant genetic structure between sites (overall F_{ST}= 0.015, $p < 0.001$). Despite the short pelagic larval duration of this species, geographic distance does not appear to correlate with genetic distances between sites. From the results gathered here, it seems possible that genetic structure at this spatial scale is driven to some extent by local population dynamics (deviations from random mating and/or a large proportion of larvae settling in proximity of relatives), yet infrequent long-distance dispersal events might also be responsible for the relatively weak spatial heterogeneity between sites. Overall, our results both highlight the utility of this new marker set for population genetic studies of this species and provide new evidence regarding the complexity of the small-scale population structure of this species.

Keywords Population genetics · Ascidiacea · Isolation by distance · Relatedness

Communicated by K. Kocot

Electronic supplementary material The online version of this article (doi:10.1007/s12526-017-0672-9) contains supplementary material, which is available to authorized users.

✉ Pablo Saenz-Agudelo
Pablo.saenzagudelo@gmail.com

[1] Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Valdivia, Chile

Introduction

The advent of next-generation sequencing has enhanced our ability to develop and characterize molecular markers (such as microsatellites or single-nucleotide polymorphisms [SNPs]) for non-model species (Guichoux et al. 2011). Hypervariable markers, such as microsatellites, exhibit rapid mutation rates and can accumulate new mutations in a much shorter time than other markers (Guichoux et al. 2011). Thus, highly variable markers are favored over more slowly evolving nuclear or mitochondrial markers when measuring recent demographic events, detecting minimal genetic variability, reconstructing pedigrees, identifying clones, and examining parentage (Selkoe and Toonen 2006).

Pyura chilensis Molina 1782, commonly known as *piure*, is a solitary tunicate (Pyuridae) endemic to the southeastern Pacific coast of South America, found from Peru (10°S) to the tenth region of Chile (44°S) (Vásquez 1983; Lancellotti and Vasquez 2000). *Pyura chilensis* inhabits subtidal rocky shores, where it can form vast patches, creating a habitat

Published online: 14 March 2017

matrix for a wide variety of organisms (Zamorano and Moreno 1975; Sepulveda et al. 2003; Sepúlveda et al. 2014). *Pyura chilensis* is a digonic, protandrous hermaphrodite, and though it has the ability to self-fertilize, it favors cross-fertilization when congeners are nearby (Manriquez and Castilla 2005). Similar to other benthic tunicates (e.g., see Ben-Shlomo et al. 2010), *P. chilensis* produces short-lived free-swimming larvae that spend 12 to 24 h in the water column before settling (Cea 1973). This short pelagic larval duration (PLD) most likely limits the dispersal potential of *P. chilensis*. Despite this, populations of other ascidians, such as *Symplegma rubra* and *Ciona intestinalis*, have been found to be highly connected across large geographic scales (Dias et al. 2006; Zhan et al. 2010); human-mediated transport has long been thought to facilitate long-range dispersal of ascidians with naturally low dispersal potential (Lambert 2007).

In addition to the ecological importance of *P. chilensis* as a provider of habitat for many species, it is an important target in Chilean coastal fisheries (Castilla and Fernandez 1998) and is an important food source for another highly valued fished abalone, the loco (*Concholepas concholepas*). Information on larval dispersal and population connectivity is critical for the effective management of both of these resources. Others have shown that knowledge of the spatial scale of dispersal can determine which conservation strategies are plausible and how reserves should be spatially arranged to produce sustainable fisheries (Russ 2002, Palumbi 2004, Sale et al. 2005, Jones et al. 2007).

To date, only two studies have measured the genetic structure of this species, and they used either partial sequences of nuclear and mitochondrial genes (Haye and Muñoz-Herrera 2013) or allozyme assays (Astorga and Ortiz 2006). Both studies reported significant population structure among the sampled locations, and the authors suggest that the extent of gene flow at evolutionary time scales is more important than was previously thought, given the low dispersal potential of the species. Furthermore, it is postulated that connectivity between geographically distant sites could be facilitated by maritime transport (hitchhiking). While these results shed light on the connectivity of this organism at evolutionary time scales, to our knowledge, the more recent demographic gene flow of *P. chilensis* has yet to be studied, most likely due to a lack of available molecular markers.

Here we describe the development of ten polymorphic microsatellite markers from next-generation sequencing data and test their application on samples from five sites in southern Chile that span approximately 80 km of coastline. These microsatellites are the first reported for this ascidian, and our results reveal the presence of a fine-scale genetic structure in this species. These markers will be extremely useful for further study of the genetic structure of populations of this important resource.

Materials and methods

Sample collection, DNA extraction, and sequencing

Total DNA was extracted from the siphon tissue of one *P. chilensis* individual from Los Molinos (−39.853, −73.396) using the High Pure PCR Template Preparation Kit (Roche, Germany), following the manufacturer's instructions. DNA concentration and purity were assessed using the Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific) and a DQ300 fluorometer (Hoefer, Inc., Holliston, MA, USA) available at the AUSTRAL-*omics* core facility of the Department of Science at the Austral University of Chile (www.australomics.cl). The DNA was then used to construct a shotgun library sequenced on a Roche Genome Sequencer (GS) Junior (Basel, Switzerland). Briefly, 500 ng of genomic DNA was randomly sheared via nebulization. Double-stranded adaptors were then blunt-ligated to the sheared DNA using the GS FLX Titanium Rapid Library Adaptor Kit. The library was then cleaned using magnetic beads (Agencourt AMPure XP; Beckman Coulter), and size selection of DNA fragments between 400 and 600 bp was visualized using an Agilent Bioanalyzer and the High Sensitivity DNA Analysis Kit (Agilent Technologies). The resulting library was then quantified using the DQ300 Hoefer fluorometer. Genomic shotgun library molecules were then clonally amplified via emulsion polymerase chain reaction (PCR) following the emPCR Amplification Method Manual Lib-L. After emulsion PCR, beads containing sufficient copies of clonally amplified library fragments were selected via the specified enrichment procedure and counted using a GS Junior Bead Counter (Roche, Germany) prior to sequencing. Following emulsion PCR enrichment, beads were deposited into the wells of a PicoTiterPlate device, and sequencing was performed. Image analysis, signal processing and base calling were performed using the GS Junior System software (454 Life Sciences/Roche).

Characterization of microsatellite loci

Using Geneious v6.1.8 software (Kearse et al. 2012), the resulting sequence data were filtered to eliminate low-quality reads or to trim reads at both ends when the chance of a base calling error was higher than 5%. The trimmed read files were then loaded into the MSATCOMMANDER software program (Faircloth 2008) to identify reads containing uninterrupted small tandem repeats (minimum of eight repeats) of di-, tri-, and tetranucleotides. MSATCOMMANDER was also used to design primer pairs on flanking regions of the sequences containing microsatellites (optimal annealing temperature = 60 °C, optimal length 20 nt, GC clamp, and remaining default parameters per Primer3 software; Untergasser et al. 2012). From all putative microsatellite sequences retrieved, 96 were selected based on

PCR product size (100–400 bp range), number of repeats (at least ten repeats for dinucleotides and eight repeats for tetranucleotides) and primer pair penalty (preference for low over high penalty). These 96 primer pairs were ordered and tested as described below.

Primer optimization

Pyura chilensis samples were collected from five sites along the coast of the Los Rios region of Chile (Table S1, Fig. 1). DNA was extracted from a total of 151 individuals from the five sites following the "HotSHOT" protocol of Meeker et al. (2007). The 96 selected primers were first tested on five samples (one per location) for amplification and polymorphism at different annealing temperatures (57–63 °C). PCR reactions were carried out using the KAPA2G Fast Multiplex PCR Kit (KAPA Biosystems). Reactions were conducted following the manufacturer's protocol, with some modifications. Reactions were run in a total volume of 10 μl containing 5 μl of Multiplex mix, 3 μl of ultrapure water, 1 μl of DNA, and 1 μl of primer (2 μM). PCRs were performed on a MultiGene OptiMax Thermal Cycler (Labnet International) with the following thermal cycling parameters: 3 min at 95 °C, followed by 30 cycles of 15 s at 95 °C, 30 s at Tm, 30 s at 72 °C, and a final extension of 1 min at 72 °C; Tm of each primer pair is listed in Table 1. Amplification success and polymorphism (size variation among individuals) were verified using a fragment analyzer (Advanced Analytical Technologies) available at the AUSTRAL-*omics* core-facility of the Faculty of Science at the Austral University of Chile (www.australomics.cl). Loci that showed clear unambiguous polymorphic PCR bands for all five individuals were selected and forward primers for these loci were ordered-labeled with Applied Biosystems (ABI) fluorescent dyes (6-FAM, NED, VIC, PET). These labeled primers were multiplexed in three different primer mixes. PCR reactions were then performed for these selected loci on all 151 samples. PCR products were sent to Macrogen (Korea) for fragment analysis on an ABI 3370XL DNA analyzer (Applied Biosystems) with GeneScan 500 LIZ (Applied Biosystems) internal size standard, following the manufacturer's instructions. Individual genotypes were scored using the microsatellite plugin of Geneious v.6.1.8.

Summary statistics

Summary statistics were generated to characterize the variability and reliability of the loci. Allelic frequencies, number of alleles (N_A), and observed (H_O) and expected heterozygosity (H_E) were estimated for each site using GenAlEx v6.5 software (Peakall and Smouse 2012). We tested for linkage disequilibrium (LD) among loci and deviations from Hardy–Weinberg equilibrium (HWE) at each site and for each locus using Genepop (Rousset 2008). Significant deviations from LD and HWE were determined using Markov chain permutations (10,000 dememorization steps, 1,000 batches, and 10,000 iterations per batch). Corrections for multiple testing were made using the false discovery rate (FDR, Benjamini and Hochberg 1995). Genepop was also used to calculate F_{IS} using the Weir and Cockerham estimate (Weir and Cockerham 1984). The frequency of null alleles at each locus was calculated using Geneland (Guillot et al. 2008), and clonality within the entire data set was determined in GenAlEx by measuring 100% multilocus matches.

Genetic differentiation and population structure

Genetic differentiation was determined via an analysis of molecular variance (AMOVA) using the GenAlEx program. Pairwise genetic differentiation (F_{ST}) was calculated in GenAlEx using the same AMOVA framework with 9999 permutations. Significant genetic differentiation was determined by FDR.

In order to determine the most likely number of populations (K-clusters) that the samples group into, Bayesian clustering was implemented in the program STRUCTURE (v2.3.4, Pritchard et al. 2000). An admixture model was used with the sample site given as a priori information. Allele frequencies were assumed to be correlated among populations. The model was run from $K = 1$ to $K = 10$. For each K, the model was run five times with a burn-in of 300,000 iterations and 500,000 subsequent iterations. STRUCTURE HARVESTER (Earl and vonHoldt 2012) was used to determine the most appropriate K based on the highest averaged maximum likelihood score and Evanno's delta K of the STRUCTURE model. Additionally, a model that accounts for null alleles was implemented in Geneland (Guillot et al. 2008). Sample site information was included in the spatial component of the model, and allele frequencies were assumed to be correlated. The model was run from $K = 1$ to $K = 10$. For each K, the model was run ten times with a burn-in of 100 and 100,000 subsequent iterations. The most appropriate K was determined based on the highest averaged maximum likelihood score of the model.

Isolation by distance

We evaluated the degree of isolation by distance (IBD) to investigate whether genetic distance between sites increased as a function of geographic distance. Linear distance (Euclidean distance in km) between pairs of sites was calculated from GPS coordinates using GenAlEx v5.1. A pairwise genetic distance matrix (F_{ST}) was calculated from the multilocus genotype data set. IBD was determined based on a Mantel test employing the pairwise linear genetic distance matrix F_{ST} [$F_{ST} / (1 - F_{ST})$] and the geographic distance

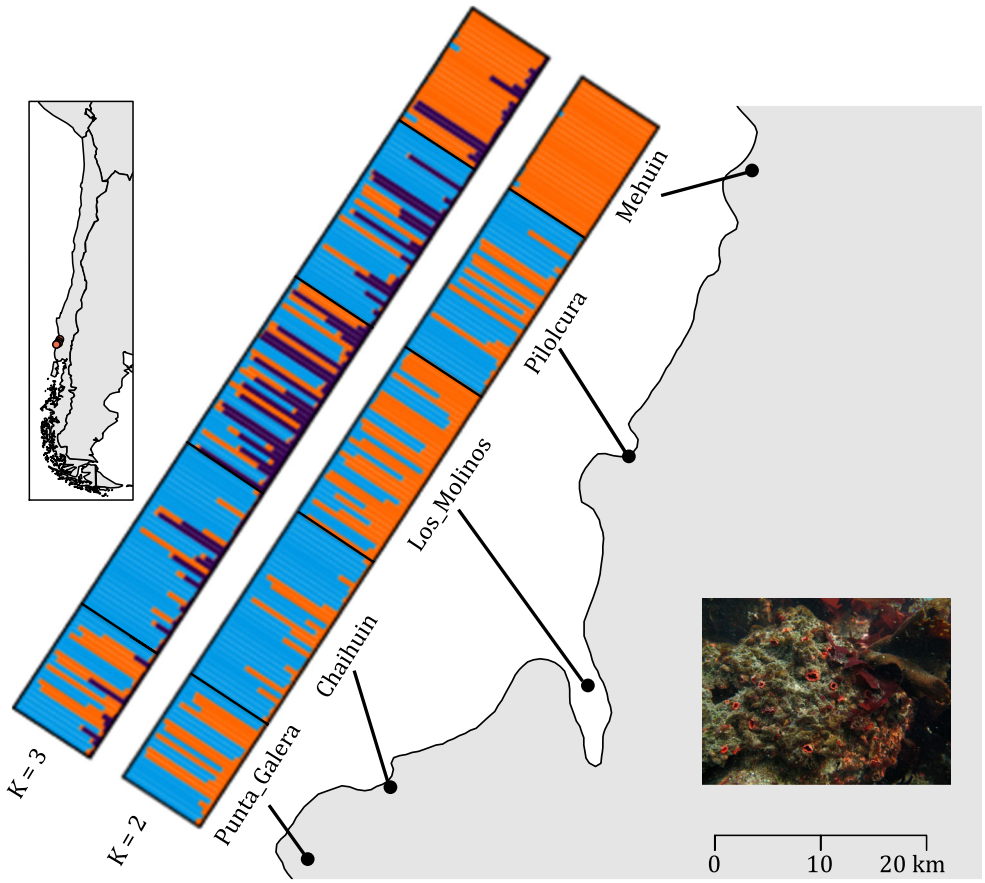

Fig. 1 Sampling sites in southern Chile (•) and photograph of *Pyura chilensis* in natural habitat near Los Molinos. Map inset shows location of sampling sites (*orange circles*) in Chile. Additionally, population genetic differentiation is shown via Bayesian clustering (STRUCTURE); the sample site was given as a priori information. Best inferred clustering was $K = 3$ determined from the mean estimated Ln(K) and $K = 2$ determined by Evanno's delta K. More information is found in Figure S1

matrix; Pearson correlation coefficients were calculated for the linear relationship between geographic distance and genetic distance, and 9999 permutations were used to test for a significant correlation between distance matrices as implemented in GenAlEx v5.1.

Spatial autocorrelation and relatedness analyses

In addition to having a short PLD, previous studies have shown that *P. chilensis* is capable of reproducing via self-fertilization (Manriquez and Castilla 2005) and that larvae prefer to settle on top of conspecifics under laboratory conditions (Manriquez and Castilla 2005). Given these observations, we analyzed genetic similarity among individuals in two different ways. First, we performed a spatial autocorrelation analysis where we tested whether genetic similarity between pairs of individuals correlated with the geographic distance between the individuals. We used the method described by Peakall and Smouse (2012) implemented in GenAlEx v5.1. Briefly, pairwise genetic distance and pairwise geographic distance matrices were generated in GenAlEx. Pairwise comparisons were then grouped into four classes according to their geographic distance. The autocorrelation coefficient (r) was then estimated for each distance class. Random permutations (10,000) and bootstrapping (10,000) were performed to test whether genotypes were distributed randomly in space ($r = 0$) or not ($r <> 0$). Significant positive r values for a given geographic distance bin would indicate that samples were more genetically similar than expected if the distribution of samples was random across space. Significant negative r values for a given geographic distance bin would indicate that samples in that distance bin were less genetically similar than expected. Under a low dispersion hypothesis, we would expect to find positive values of r for samples that were geographically close together, and zero or negative values of r for samples that were geographically distant. Second, we tested whether average relatedness within sites was higher than the average relatedness across the entire data set. Significant mean within-site relatedness would also support the idea of strong local population dynamics in natural populations of this species. For this, we calculated pairwise genetic relatedness in GenAlEx v5.1 using the Lynch and Ritland (1999) estimator. We chose this estimator because it has been suggested to be particularly useful in the case of multiple hypervariable loci (Lynch and Ritland 1999). We then computed average relatedness within each site and

Table 1 Microsatellite markers isolated from *Pyura chilensis* and amplified in samples from five sites: Mehuín, Pilolcura, Los Molinos, Chaihuín, and Punta Galera. Annealing temperature (*Ta*), number of individuals amplified (*N*), number of alleles (*Na*), observed heterozygosity (*Ho*), unbiased expected heterozygosity (*uHe*), and inbreeding coefficient (F_{IS}, Genepop W&C) are given for each locus and each site. Significant *p* values after correction by false discovery rate (alpha = 0.05) are shown in bold and italics

Locus	Primer sequence F and R (5′–3′)	Ta (°C)	Repeat motif	Size range (bp)	Mehuín						Pilolcura					
					N	*Na*	*Ho*	*uHe*	*Fis*	*p* value	*N*	*Na*	*Ho*	*uHe*	*Fis*	*p* value
PIU76	L:GTTTATGTCACCGGCCTACTTCC R:ACAGCAGTTGATTCACAGCC	57	AG(14)	342–474	12	11	0.167	0.913	0.824	***0***	31	18	0.613	0.915	0.3341	***0***
PIU66	L:ATTAGTCAGACAGGCAGGCC R:GTTTCACAGCATACACAGTGCG	57	AG(14)	232–298	24	10	0.625	0.736	0.1534	0.1407	34	12	0.500	0.770	0.3541	***0.0001***
PIU20	L:GTTTCAGGAACGGACGAGATCTC R:CGGCTGTGCTTGCATAATTC	57	AGAT(10)	202–298	23	16	0.913	0.926	0.0139	0.8193	33	16	0.818	0.908	0.1005	***0.0332***
PIU67	L:GACGAGCGTTGTGAGAAAGG R:GTTTAAAGACTTCAACGACCGCTC	57	AC(14)	210–254	24	11	0.833	0.873	0.0466	0.0274	31	12	0.742	0.906	0.1839	***0.0121***
PIU19	L:GGATCAGCAGCAACGTAGAC R:GTTTATTGTCACGGCTGCCAC	57	AGAT(11)	267–379	23	16	0.739	0.944	0.2208	0.04	32	14	0.750	0.903	0.172	***0.0325***
PIU90	L:AGGTAGCCCTTCGCTTTGAT R:GTCAGGGTGTCGTAATTGGG	57	AG(12)	200–226	24	11	0.708	0.895	0.2125	0.0958	34	12	0.824	0.875	0.0595	0.5215
PIU36	L:GTTTGAATACATACCACGGCTGC R:GAGTCCGAGTACATGGGTGG	57	AGAT(11)	288–392	12	9	0.250	0.866	0.7203	***0***	19	10	0.421	0.838	0.5043	***0.0004***
PIU82	L:TACCTTTGACCTCAGGGAGC R:TATATGCCGCAAGTGTGGTC	57	AGAT(11)	116–296	24	12	0.583	0.769	0.245	***0.0032***	33	17	0.667	0.898	0.2605	***0.0027***
PIU17	L:GTTTGGTCGTATCGTCACAACCC R:TCGAGGCTCCTAAATTCCCG	57	AGAT(10)	156–296	16	12	0.500	0.925	0.4678	***0***	34	14	0.559	0.910	0.3892	***0***
PIU06	L:AGAAGGAGAATGGTGCGTCC R:GTTTAAGGATGAGCACACGAGTAC	57	AT(10)	173–195	23	11	0.696	0.830	0.1649	0.2374	33	8	0.909	0.672	−0.3607	***0***

Locus	Los Molinos						Chaihuín						Punta Galera					
	N	*Na*	*Ho*	*uHe*	*Fis*	*p* value	*N*	*Na*	*Ho*	*uHe*	*Fis*	*p* value	*N*	*Na*	*Ho*	*uHe*	*Fis*	*p* value
PIU76	31	20.000	0.484	0.939	0.4886	***0***	32	16.000	0.688	0.879	0.2206	***0.0149***	18	10.000	0.500	0.870	0.4323	***0.0003***
PIU66	36	9.000	0.583	0.723	0.1954	***0.0001***	35	9.000	0.571	0.779	0.2696	***0.008***	22	12.000	0.636	0.814	0.2222	***0.0114***
PIU20	36	16.000	0.944	0.934	−0.011	0.3205	35	16.000	0.600	0.891	0.3299	0	22	11.000	0.773	0.910	0.154	0.0613
PIU67	33	10.000	0.636	0.807	0.2136	0.0527	29	10.000	0.724	0.834	0.134	0.2285	19	11.000	0.842	0.859	0.0204	0.8028
PIU19	35	21.000	0.714	0.942	0.2444	***0***	35	16.000	0.800	0.916	0.1286	***0.1436***	22	14.000	0.864	0.919	0.0612	0.2818
PIU90	36	12.000	0.861	0.873	0.0141	0.4025	35	11.000	0.829	0.881	0.0605	0.2513	22	10.000	0.591	0.855	0.3141	***0.0016***
PIU36	20	11.000	0.350	0.828	0.5837	***0***	31	11.000	0.387	0.882	0.565	***0***	15	8.000	0.200	0.839	0.768	***0***
PIU82	35	21.000	0.543	0.884	0.3891	***0***	35	16.000	0.629	0.855	0.2674	***0***	21	16.000	0.714	0.909	0.2188	***0.0032***
PIU17	32	18.000	0.563	0.924	0.3948	***0***	34	20.000	0.735	0.951	0.2293	***0.0001***	17	16.000	0.647	0.927	0.3084	***0.0008***

performed 10,000 permutations to test whether mean within-site relatedness was significantly different from the overall mean relatedness across all sites.

Results

Sequencing and characterization of microsatellite loci

A total of 233,858 reads were generated from the sequenced 454 library. Of these, 146,389 reads passed all quality and length filter controls. Thus, the final data set used to mine microsatellites included 146,389 reads, or 57,203,311 bases, with an average length of 391 bases. From this, 2323 microsatellites were recovered; primers were designed and ordered for 96 microsatellites.

Primer optimization and summary statistics

Of the 96 primers pairs that were designed and tested, only ten were polymorphic and could be readily amplified and genotyped. The 454 sequence reads for the ten microsatellites presented here can be found in GenBank (accession nos. KX775329–KX775338). The average number of alleles per locus (*Na*) ranged from 9.8 to 16.4; see Table 1 for main summary statistics. The average observed heterozygosity across sites (*Ho*) ranged from 0.322 to 0.810, while the average unbiased expected heterozygosity across sites (*uHe*) ranged from 0.714 to 0.927. We found that all loci showed significant deviations from HWE in at least one site after FDR corrections were made (34 significant tests after FDR, of 50 tests performed); from these, four loci displayed consistent significant deviations from HWE in all sites (PIU76, PIU36, PIU82, and PIU17; see Table 1 for details). Given the biology of this species, it is possible that these deviations from HWE are due to non-random mating within sites and not to the presence of null alleles. However, we note that of the four loci with significant HWE deviations in all sites, two (PIU76 and PIU36) displayed >10% missing data, which is consistent with the hypothesis of the presence of null alleles. In addition, we found that of 225 pairwise tests, 25 (11%) exhibited significant deviations from linkage equilibrium (after FDR correction). Therefore, 17 pairwise comparisons were significant at Los Molinos, one at Mehuín, and seven at Pilolcura. However, a global comparison of all linkage disequilibrium tests across sites revealed that no pairwise comparisons of loci were significantly consistent at all sites. Lastly, none of the samples genotyped had 100% multilocus match-ups, indicating a lack of clones in the data set.

Genetic differentiation and population structure

The AMOVA indicated that the partition of genetic variation among sites was low ($F_{ST} = 0.015$) but significant ($p = 0.0001$). Additionally, the variation among samples within sites was moderately high ($F_{IS} = 0.200$) and significant ($p = 0.0001$), suggesting departures from random mating within sites. It should be noted, however, that the high and significant degree of within-individual variation ($F_{IT} = 0.213$, $p = 0.0001$) that is very common for highly variable markers such as microsatellites likely reduces the ability to detect variation among and within sites. Despite the low overall variability detected in the global AMOVA, all of the pairwise comparisons of genetic differentiation among sites were significant even after FDR correction (Table 2); F_{ST} values ranged from 0.005 to 0.035. Overall, the highest and most significant F_{ST} values involved pairwise comparisons with Mehuín.

The results of the STRUCTURE analysis indicated that the data likely cluster in two to three independent groups depending on the summary statistic used to choose the most likely number of clusters (Fig. 1). When using the natural log of the maximum likelihood score, $K = 3$ was the most probable number of clusters that the data grouped into. Alternatively, Evanno's delta K indicated that $K = 2$ was the most probable number of data clusters (Figure S1). Both STRUCTURE bar plots are included in Fig. 1. Interestingly, both clustering solutions clearly separate the Mehuín site from the rest and suggest potential uniqueness of Los Molinos; however, some degree of admixture was detected among all sites. It is also worth noting that the results from the Bayesian clustering model implemented in Geneland, which accounts for the presence of null alleles, indicated that the data most likely partition into three clusters (Table S1). These results are largely in agreement with those from STRUCTURE. According to Geneland, Mehuín samples grouped together in a single cluster, a second cluster contained the samples from Los Molinos, and a third cluster included all remaining samples (Figure S1).

Due to the estimated high frequency of null alleles in two of the loci, PIU76 and PIU36, these loci were removed from the data set, and STRUCTURE and Geneland were rerun. The removal of these loci had a minimal effect on the STRUCTURE and Geneland results. The most probable number of data clusters for the eight-locus data set included $K = 3$ or $K = 5$ (STRUCTURE, natural log of the maximum likelihood), $K = 2$ (STRUCTURE, Evanno's delta K), and $K = 3$ (Geneland). Regardless of the model used, and with both ten and eight loci, the data indicate the clear separation of the Mehuín site. In all $K = 3$ models, Mehuín samples cluster independently, Los Molinos samples cluster independently, and

Table 2 Pairwise population differentiation based on ten microsatellite loci. Sites are indicated from north to south. F_{ST} values are below the diagonal, and p values are above the diagonal. All values were significant after FDR correction (alpha = 0.05). Values were calculated with 9999 permutations of the full data set

	Mehuín	Pilolcura	Los Molinos	Chaihuín	Pta. Galera
Mehuín		0.000	0.000	0.000	0.000
Pilolcura	0.029		0.000	0.031	0.035
Los Molinos	0.021	0.013		0.001	0.001
Chaihuín	0.035	0.005	0.011		0.027
Pta. Galera	0.022	0.006	0.012	0.007	

the remaining sites cluster together. In the $K = 5$ model, all sites cluster independently of one another.

Isolation by distance

The IBD Mantel test indicated that the relationship between geographic distance and linearized genetic distance was not significant (Fig. 2, $p = 0.137$). Nevertheless, there did appear to be a trend that as geographic distance increased, so did genetic distance ($R^2 = 0.29$).

Spatial autocorrelation and relatedness analyses

We found that spatial autocorrelation was positive and significant for sample comparisons within sites ($r = 0.01$, $p < 0.001$) but was not different from zero for samples that were separated by up to 60 km. The spatial correlation coefficient was negative and significantly different from zero for all sample pairs that were separated by 60 km or more ($r = -0.023$, $p < 0.001$; Fig. 3a).

Overall, mean within-site relatedness ranged from −0.003 to 0.019. Interestingly, this was significantly higher than mean relatedness across all sites (Fig. 3b). The mean relatedness of Mehuín samples was the highest compared to all other sites ($r = 0.019$, $p < 0.001$). Similar results were obtained when these analyses were performed excluding loci PIU76 and PIU36 (data not shown).

Discussion

Here, we have presented ten polymorphic microsatellite loci for use in characterizing the demographic connectivity of *Pyura chilensis*, a commercially and ecologically relevant ascidian of the southeastern Pacific. Additionally, we have shown for the first time that populations of *P. chilensis* can be structured at scales of tens of kilometers, though geographic distance does not appear to be the main determinant of such structure. Interestingly, we also found higher relatedness and autocorrelation between genetic distance and geographic distance within sites, which is in agreement with what is known about the reproductive biology of the species and its weak larval dispersal potential. Thus it is possible that the patterns of genetic structure detected may indicate that individuals within sites are more closely related than would be expected for a randomly mating population, yet infrequent long-distance dispersal events minimize detectable spatial heterogeneity between sites. The markers developed here will be invaluable tools with which to gain a better understanding of the recent demographic history and potential barriers to connectivity of this species, information that will no doubt be useful in local species management.

The results from this study show that *P. chilensis* populations can be structured at a scale of tens of kilometers. Specifically, genetic structure was found in *P. chilensis* samples from sites separated by a minimum of 10 km and a maximum of 80 km, depending on the pairwise site comparison. The results of the AMOVA, STRUCTURE, and Geneland analyses conclusively illustrate that there are mechanisms at

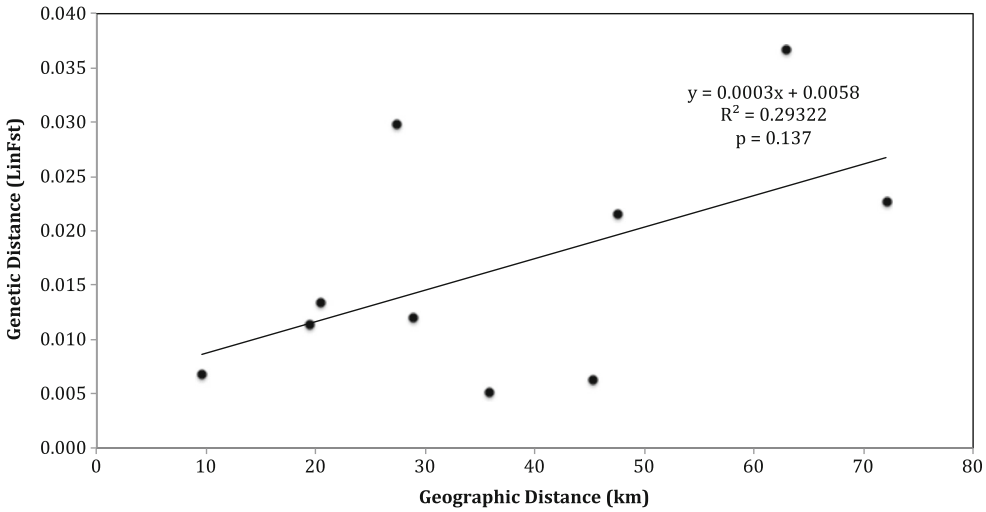

Fig. 2 IBD analysis. The Mantel test scatterplot of IBD shows genetic distance [$F_{ST} / (1 - F_{ST})$] as a function of geographic distance. No significant correlation was found

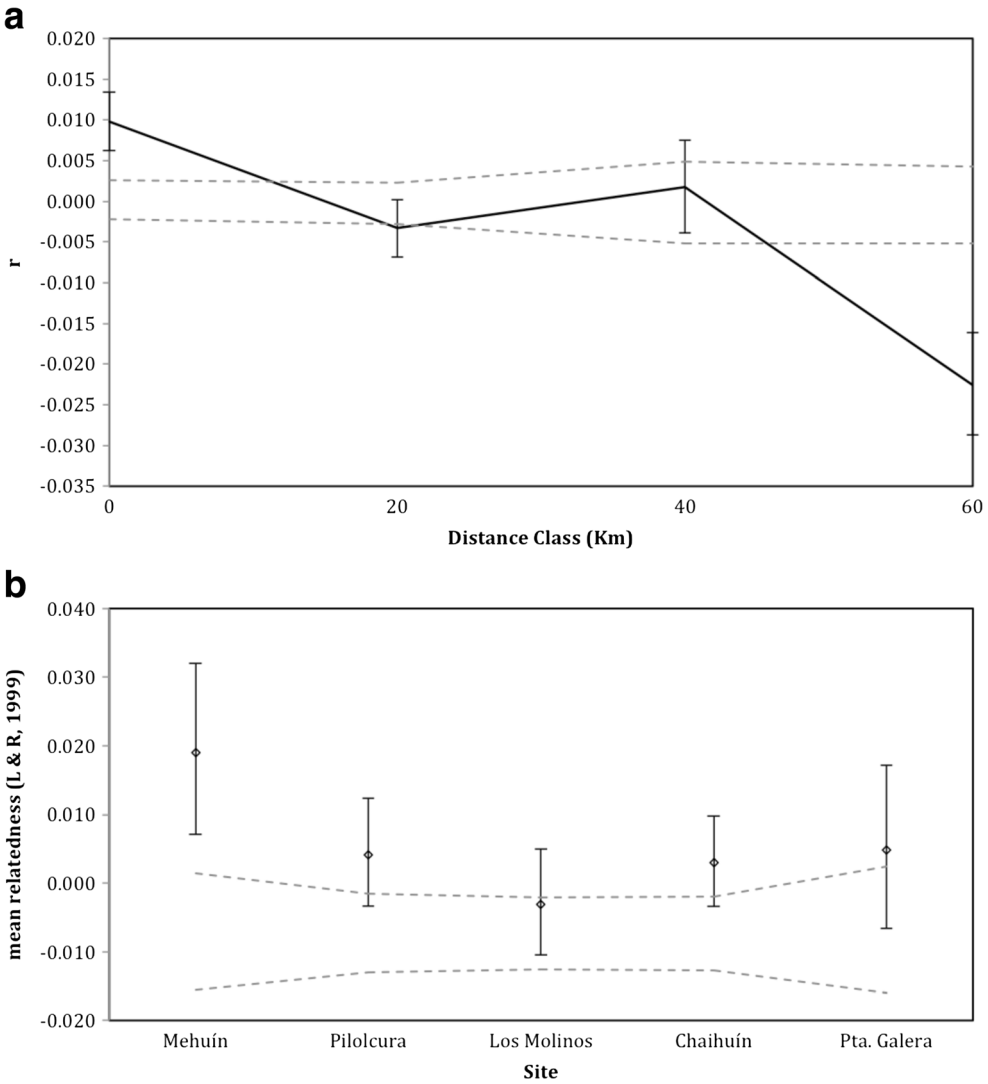

Fig. 3 **a** Spatial autocorrelation results. Average spatial autocorrelation is indicated as a *solid black line* that connects all distance class bins. *Error bars* determine the 95% confidence intervals around *r* as determined by 10,000-bootstrap resampling. *Gray dashed lines* indicate 95% confidence intervals around the null hypothesis of no spatial structure for the combined data set. **b** Lynch and Ritland (1999) estimates of mean relatedness. *Diamonds* represent within-site mean relatedness, and *error bars* indicate 95% confidence intervals around the mean calculated with bootstrapping. The *gray dashed lines* indicate the 95% confidence intervals around the null hypothesis of no difference across sites

this spatial scale that prevent complete genetic homogenization of *P. chilensis* populations. Interestingly, considering the benthic nature and perceived low dispersal potential of *P. chilensis*, the genetic variability found here was not significantly correlated with geographic distance. This was surprising considering that *P. chilensis* larvae are thought to remain in the water column for no more than 24 h (Cea 1969). Our result is similar to that found by Haye and Muñoz-Herrera (2013) using mitochondrial and nuclear markers at a much larger geographic scale. They also found a lack of isolation by distance for *P. chilensis* samples spread across 1700 km of coastline. The authors hypothesized that, as has been noted for other ascidian species (Castilla et al. 2002; Nóbrega et al. 2004), long-distance dispersal could be facilitated by transportation on boat hulls or floating debris. In fact, human-mediated transport has been implicated for many ascidian species (Lambert and Lambert 2003; Lambert 2007; Ben-Shlomo et al. 2010; Zhan et al. 2010). Recent colonization by few individuals transported in ship ballast or on ship hulls would explain high within-site relatedness yet a lack of isolation by distance. Alternatively, however, gene flow at demographic scales could be facilitated or restricted by other, yet untested, natural phenomena. Recent studies have noted that environmental heterogeneity and oceanographic features play a larger role in determining population structure than does linear geographic distance (Galindo et al. 2010; Selkoe et al. 2010; Giles et al. 2015; Saenz-Agudelo et al. 2015).

Here we found both that mean genetic relatedness within sites was higher than what would be expected by chance, and there was a positive spatial autocorrelation for within-site comparisons. In addition, while previous studies of other ascidians have documented that some species are capable of clonal budding as a reproductive strategy, clonal budding does not appear to be a common reproductive strategy of *P. chilensis*, as no clones were found among the 151 samples analyzed here. While the presence of null alleles in the loci that deviated from HWE cannot be ruled out, the results were robust to the removal of loci with the greatest deviations from

HWE. Therefore, it is possible that these deviations are simply the reflection of deviations from random mating in natural *P. chilensis* populations at the spatial scale sampled (tens of meters). Previous studies have shown that *P. chilensis* is a simultaneous hermaphrodite and is capable of self-fertilization (Manriquez and Castilla 2005); thus minimal self-fertilization events or mating with congeners could lead to drastic decreases in heterozygosity. Furthermore, other population genetic studies of ascidians have also noted deviations from HWE with heterozygote deficits (for examples see Ben-Shlomo et al. 2006; Pérez-Portela and Turon 2008; Dupont et al. 2009; Ben-Shlomo et al. 2010). Our results suggest that population dynamics even at a scale of tens of kilometers, as well as mating patterns in natural populations of this species, may be more complex than previously thought. Further studies at fine spatial scales will be needed to elucidate the species' reproductive patterns.

Rare long-distance dispersal events paired with frequent self-recruitment and a lack of non-random mating at local scales could explain both the broad-scale structure and within-site patterns found here. Based on the detected genetic structure at tens of kilometers, *P. chilensis* likely often recruits locally; however, occasional long-ranging dispersal or transport of juveniles or adults among sites mediated by humans or strong tidal waves could explain, to some degree, the lack of IBD found here. Further to this, we found that in sites that were more differentiated, the mean relatedness between samples was higher. While we can only speculate about the reasons for this particular pattern, if individuals generally recruit very close to their relatives, this would facilitate reproduction among relatives over time, which in turn would explain the high genetic relatedness and autocorrelation within sites. It has been well documented in plants that repeated selfing or mating with conspecifics leads to enhanced spatial genetic structure (Vekemans and Hardy 2004).

In summary, this study characterizes novel polymorphic microsatellite loci that will be important tools for gaining a better understanding of demographic connectivity and barriers to gene flow in *Pyura chilensis*. Future studies involving increased sampling at multiple scales will be necessary to fully elucidate the dominant mode of reproduction and genetic diversity of this species. The information provided here, and that which will arise from future studies involving these microsatellite markers, will no doubt be valuable for the effective management of this commercial ascidian. In other systems, the detection of genetic structure at hundreds of meters to tens of kilometers has provided support for fisheries management at a local scale (Costantini et al. 2007; Almany et al. 2013). As it does not appear that *P. chilensis* is genetically homogeneous along the southern coast of Chile, multiple local rather than few large reserves could help protect genetic diversity distributed at small scales. However, additional studies are needed to determine the factors contributing to the population structure of *P. chilensis* in order to inform management plans.

Acknowledgements This study was supported by CONICYT, programa FONDECYT grant no. 11140121 to P.S-A. All sampling was performed with authorization from the Subsecretaria de Pesca y Acuicultura de Chile (permit: P.INV 241/2014). In addition, we thank The Nature Conservancy, Marcelo Antillanca, the Sindicato de Pescadores de Chaihuín, and Jose Martel from Los Molinos for facilitating sample collection in Chaihuín and Los Molinos.

References

Almany GR, Hamilton RJ, Bode M, Matawai M, Potuku T, Saenz-Agudelo P, Planes S, Berumen ML, Rhodes KL, Thorrold SR, Russ GR, Jones GP (2013) Dispersal of grouper larvae drives local resource sharing in a coral reef fishery. Curr Biol 23:626–630

Astorga MP, Ortiz JC (2006) Variabilidad genética y estructura poblacional del tunicado *Pyura chilensis* Molina, 1782, en la costa de Chile. Rev Chil Hist Nat 79:423–434. doi:10.4067/S0716-078X2006000400002

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 57:289–300

Ben-Shlomo R, Paz G, Rinkevich B (2006) Postglacial-period and recent invasions shape the population genetics of Botryllid ascidians along European Atlantic coasts. Ecosystems 9:1118–1127. doi:10.1007/s10021-006-0141-y

Ben-Shlomo R, Reem E, Douek J, Rinkevich B (2010) Population genetics of the invasive ascidian *Botryllus schlosseri* from Southern American coasts. Mar Ecol Prog Ser 412:85–92. doi:10.3354/meps08688

Castilla JC, Fernandez M (1998) Small-scale benthic fisheries in Chile: on co-management and sustainable use of benthic invertebrates. Ecol Appl 8:S124–S132. doi:10.1890/1051-0761(1998)8[S124:SBFICO]2.0.CO;2

Castilla JC, Collins AG, Meyer CP, Guíñez R, Lindberg DR (2002) Recent introduction of the dominant tunicate, *Pyura praeputialis* (Urochordata, Pyuridae) to Antofagasta, Chile. Mol Ecol 11:1579–1584

Cea G (1969) Estadios primarios de desarrollo y metamorfosis de *Pyura chilensis* Molina, 1782 (Tunicata, Ascidiacea, Pyuridae). Bol Soc Biol Concepción, Tomo XLII:317–331

Cea G (1973) Biología del Piure (*Pyura chilensis* Molina, 1782; Chordata, Tunicata, Ascidiacea). Gayana Zool 28:1–65

Costantini F, Fauvelot C, Abbiati M (2007) Fine-scale genetic structuring in *Corallium rubrum*: evidence of inbreeding and limited effective larval dispersal. Mar Ecol Prog Ser 340:109–119

Dias GM, Duarte LFL, Solferini VN (2006) Low genetic differentiation between isolated populations of the colonial ascidian *Symplegma rubra* Monniot, C. 1972. Mar Biol 148:807–815

Dupont L, Viard F, Dowell MJ, Wood C, Bishop JDD (2009) Fine- and regional-scale genetic structure of the exotic ascidian *Styela clava* (Tunicata) in southwest England, 50 years after its introduction. Mol Ecol 18:442–453. doi:10.1111/j.1365-294X.2008.04045.x

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. doi:10.1007/s12686-011-9548-7

Faircloth BC (2008) Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol Ecol Resour 8:92–4. doi:10.1111/j.1471-8286.2007.01884.x

Galindo HM, Pfeiffer-Herbert AS, McManus MA, Chao Y, Chai F, Palumbi SR (2010) Seascape genetics along a steep cline: using genetic patterns to test predictions of marine larval dispersal. Mol Ecol 19:3692–3707. doi:10.1111/j.1365-294X.2010.04694.x

Giles EC, Saenz-Agudelo P, Hussey NE, Ravasi T, Berumen ML (2015) Exploring seascape genetics and kinship in the reef spong *Stylissa carteri* in the Red Sea. Ecol Evol 5:2487–2502. doi:10.1002/ece3.1511

Guichoux E, Lagache L, Wagner S et al (2011) Current trends in microsatellite genotyping. Mol Ecol Resour 11:591–611. doi:10.1111/j.1755-0998.2011.03014.x

Guillot G, Santos F, Estoup A (2008) Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. Bioinformatics 24:1406–1407. doi:10.1093/bioinformatics/btn136

Haye PA, Muñoz-Herrera NC (2013) Isolation with differentiation followed by expansion with admixture in the tunicate *Pyura chilensis*. BMC Evol Biol 13:252. doi:10.1186/1471-2148-13-252

Jones GP, Srinivasan M, Almany GR (2007) Population connectivity and conservation of marine biodiversity. Oceanography 20:100–111

Kearse M, Moir R, Wilson A et al (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–9. doi:10.1093/bioinformatics/bts199

Lambert G (2007) Invasive sea squirts: a growing global problem. J Exp Mar Biol Ecol 342(1):3–4

Lambert CC, Lambert G (2003) Persistence and differential distribution of nonindigenous ascidians in harbors of the Southern California Bight. Mar Ecol Prog Ser 259:145–161

Lancellotti DA, Vasquez JA (2000) Zoogeografía de macroinvertebrados bentónicos de la costa de Chile: contribución para la conservación marina. Rev Chil Hist Nat 73:99–129. doi:10.4067/S0716-078X2000000100011

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. Genetics 152(4):1753–1766

Manriquez PH, Castilla JC (2005) Self-fertilization as an alternative mode of reproduction in the solitary tunicate *Pyura chilensis*. Mar Ecol Prog Ser 305:113–125. doi:10.3354/meps305113

Meeker ND, Hutchinson SA, Ho L, Trede NS (2007) Method for isolation of PCR-ready genomic DNA from zebrafish tissues. Biotechniques 43:610–614. doi:10.2144/000112619

Nóbrega R, Solé-Cava AM, Russo CAM (2004) High genetic homogeneity of an intertidal marine invertebrate along 8000 km of the atlantic coast of the Americas. Journal of Experimental Marine Biology and Ecology 303(2):173-181

Palumbi SR (2004) Marine reserves and ocean neighborhoods: the spatial scale of marine populations and their management. Annu Rev Env Resour 29:31–68

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537–9. doi:10.1093/bioinformatics/bts460

Pérez-Portela R, Turon X (2008) Cryptic divergence and strong population structure in the colonial invertebrate *Pycnoclavella communis* (Ascidiacea) inferred from molecular data. Zoology 111(2):163–178

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2):945–959

Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour 8:103–6. doi:10.1111/j.1471-8286.2007.01931.x

Russ GR (2002) Yet another review of marine reserves as reef fishery management tools. In Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem, PF Sale, ed. 421–443

Saenz-Agudelo P, Dibattista JD, Piatek MJ, Gaither MR, Harrison HB, Nanninga GB, Bermuen ML (2015) Seascape genetics along environmental gradients in the Arabian Peninsula: insights from ddRAD sequencing of anemonefishes. Mol Ecol 24:6241–6255, **10.111/mec.13471**

Sale PF, Cowen RK, Danilowicz BS, Jones G, Kritzer J, Lindeman K, Planes S, Polunin N, Russ G, Sadovy Y (2005) Critical science gaps impede use of no-take fishery reserves. Trends Ecol Evol 20:74–80. doi:10.1016/j.tree.2004.11.007

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol Lett 9:615–629. doi:10.1111/j.1461-0248.2006.00889.x

Selkoe KA, Watson JR, White C, Horin TB, Iacchei M, Mitarai S, Siegel DA, Gaines SD, Toonen RJ (2010) Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. Mol Ecol. doi:10.1111/j.1365-294X.2010.04658.x

Sepulveda R, Cancino JM, Thiel M (2003) The peracarid epifauna associated with the ascidian *Pyura chilensis* (Molina, 1782) (Ascidiacea: Pyuridae). J Nat Hist 37:1555–1569. doi:10.1080/00222930110099615

Sepúlveda RD, Rozbaczylo N, Ibáñez CM et al (2014) Ascidian-associated polychaetes: ecological implications of aggregation size and tube-building chaetopterids on assemblage structure in the Southeastern Pacific Ocean. Mar Biodivers 45:733–741. doi:10.1007/s12526-014-0283-7

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3 - new capabilities and interfaces. Nucleic Acids Res 40(15):e115

Vásquez J (1983) *Pyura chilensis* Molina 1782 en el Norte del Perú (Ascidacea, Pyuridae). Bol Soc Biol Concepc 54:171–172

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. Mol Ecol 13:921–935

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 1358–1370

Zamorano J, Moreno C (1975) Comunidades bentónicas del sublitoral rocoso de Bahía Corral, 1: Area mínima de muestreo y descripción cuantitativa de la asociación de *Pyura chilensis* Molina. Medio Ambient 1:58–66

Zhan A, Macisaac HJ, Cristescu ME (2010) Invasion genetics of the Ciona intestinalis species complex: from regional endemism to global homogeneity. Mol Ecol 19(21):4678–4694