Molecular Phylogenetics and Evolution 101 (2016) 19–31

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Population genetics features for persistent, but transient, *Botryllus schlosseri* (Urochordata) congregations in a central Californian marina

Arzu Karahan [a,b,*], Jacob Douek [b], Guy Paz [b], Baruch Rinkevich [b]

[a] Middle East Technical University, Institute of Marine Science, Department of Marine Biology and Fisheries, Mersin, Turkey
[b] Israel Oceanographic and Limnology Research, National Institute of Oceanography, Tel Shikmona, PO Box 8030, Haifa 31080, Israel

ARTICLE INFO

Article history:
Received 9 February 2015
Revised 1 May 2016
Accepted 2 May 2016
Available online 3 May 2016

Keywords:
Ascidian
COI
Heterozygote deficiency
Microsatellite
Moss Landing
Phylogeography

ABSTRACT

The colonial tunicate *Botryllus schlosseri* is a globally distributed, invasive ascidian that has colonized the Californian coasts of the USA during the mid-late 1940s and has, since the late 1980s, spread north to Washington. This study analyzes the population genetic characteristics of transient populations residing at the Elkhorn Yacht-Club (EYC), in central California (seven sessions, 1996–2008), which suffered periodic catastrophes caused by episodic fresh-water floods and a single sampling session (in the year 2001) of five West-Coast populations using the mtDNA COI gene and five microsatellite markers. EYC microsatellite results were further compared with the closely situated but persistent population of the Santa Cruz Harbor (SCH) to understand the impact on EYC population regeneration processes after the 2005-flood catastrophe. All microsatellites were highly polymorphic, revealing a large number of unique alleles at different sampling dates. Whereas pairwise θ did not reveal significant differences between the EYC time-series samplings, the overall θ was significant, as it was between all the 2001 West Coast populations. The most likely cluster number was 3 for the EYC samples whereas two *K* values were obtained (2 and 5) for the 2001 samples. Tajima's *D* and Fu's/*Fs* tests did not reject the null hypothesis for COI neutral evolution, except for in the EYC-2000, 2007 and two 2001 samplings. The wide geographical range of the analyses has indicated that following the EYC 2005-flood catastrophe, newcomers could have originated from neighboring populations, from deep-water colonies that may have escaped the 2005 low salinity event, or less expectedly, from far away West-Coast populations, while revealing that the SCH population is the most probable source for the EYC population.

1. Introduction

The colonial tunicate *Botryllus schlosseri* is a globally distributed ascidian, a common member of shallow water (<200 m depth), marine hard-bottom sessile assemblages in the northern and southern hemispheres (Ruiz et al., 2000; Lambert, 2001; Stoner et al., 2002; Paz et al., 2003; Ben-Shlomo et al., 2006, 2010; Bock et al., 2012; Reem et al., 2013a, 2013b). In the wild, adults settle on natural and manmade structures (primarily in marinas and harbors), on/under stones, on algae and seaweed, floats, dock pilings, ships' hulls and other artificial substrata. The planktonic phase of *B. schlosseri* larvae is highly restricted (<1 h), and therefore the poorly swimming larvae settle in close proximity to their parental colonies (Grosberg, 1987; Rinkevich and Weissman, 1987), limiting population dispersal to local sites. Therefore, the extensive expansion of *B. schlosseri* colonies worldwide into so many disparate temperate and sub-tropical marine habitats has been based on the ability of adult colonies to withstand long journeys via attachment to moving substrates such as ship hulls, floating objects and the surfaces of commercially important organisms, mainly crabs and oysters, that are transported across oceans (Paz et al., 2003; Bernier et al., 2009).

B. schlosseri colonies are found in all of the Mediterranean Sea as well as in European waters, extending as far north as Norway, and also are distributed throughout Asia (Korea, Japan, Hong Kong and India), Australia, Tasmania, New Zealand, South Africa, and the east and West Coasts of Northern and Southern America (Ruiz et al., 2000; Stoner et al., 2002; Paz et al., 2003; Ben-Shlomo et al., 2006, 2008, 2010; Reem et al., 2013a, 2013b). While the USA east coast populations originated from ancient Mediterranean/Atlantic introductions (the first Massachusetts records date back to 1841; Ruiz et al., 2000), the first USA West Coast introductions were recorded in central California during the mid-to-late 1940s (USA Navy, 1951; Kozloff, 1975), presumably originating from Western Pacific populations (Lejeusne et al., 2011).

* Corresponding author at: Middle East Technical University, Institute of Marine Science, Department of Biology, Mersin, Turkey.
E-mail addresses: arzukarahan@ims.metu.edu.tr (A. Karahan), douek@ocean.org.il (J. Douek), guy@ocean.org.il (G. Paz), buki@ocean.org.il (B. Rinkevich).

http://dx.doi.org/10.1016/j.ympev.2016.05.005
1055-7903/

Under natural conditions, *B. schlosseri* colonies are short-lived organisms, with a seasonally dependent life-span of 82–247 days recorded for the USA West Coast populations (Chadwick-Furman and Weissman, 1995), a significantly reduced life expectancy compared to laboratory maintained genotypes, some of which exhibited extended life-spans of 7–13 years (Rinkevich and Shapira, 1998; Rinkevich et al., 2013). As *B. schlosseri* populations typically produce about five generations/year (Chadwick-Furman and Weissman, 1995), in the decades since the tunicate's introduction to the coastal sites of central California it has reproduced hundreds of successive sexual generations, underlying the successful establishment of an invasive species with populations in equilibrium (Reem et al., 2013a). *B. schlosseri* colonies from Monterey Bay, central California, attain sexual maturity within 49 days following settlement, and achieve colonial sizes of up to 1400 zooids within 69 days (Chadwick-Furman and Weissman, 1995), all attesting to the potential rapid turnover in population genetic characteristics during the seven decades since its arrival in central California.

While the literature assesses at length numerous cases of population genetic structures in marine organisms over assorted geographic scales, an insufficient number of studies endeavor to evaluate population genetics' alternations within a single natural population over an extended scale of generations (e.g., Pérez-Portela et al., 2012; Reem et al., 2013a). Recently, Reem et al. (2013a) have studied a single, relatively isolated population of *B. schlosseri* from the central Californian coast (Santa Cruz Harbor; SCH) that was sampled over a period of 13 years (nine session, 1995–2008). Alleles found at five microsatellite loci revealed that the SCH *B. schlosseri* population had been exposed to a low frequency of repeated invasions for prolonged periods, supporting the notion of relative genetic isolation, and had been influenced by genetic drift and high mutation rates, causing the appearance of numerous newly added microsatellite alleles and the loss of established alleles (Reem et al., 2013a). This status for population genetic structures is not ubiquitous to the tunicates as revealed for some pelagic tunicates (Lavaniegos and Ohman, 2003).

The Elkhorn Yacht Club (EYC) harbor lies about 27 km south-east of SCH, near the openings of Elkhorn Slough (the second largest marine wetland in California), Moro Cojo Slough and the Salinas River, and is thus subjected to episodic freshwater flooding events that occur once every several years (CalAm Coastal Water Project Final Environmental Impact Report; http://www.water.ca.gov/sfmp/resources/Attachment_C_History_Appendices_A-F.pdf). More severe than usual, these floods caused the complete eradication of the shallow water marine assemblages (including all botryllid ascidians) in the area for several months, until the marine biota reestablished itself over the following spring–summer months (BR pers.-observ., Fig. S1). These EYC environmental conditions enabled us to compare the temporal genetic structure dynamics of this transient *B. schlosseri* population with the geographically adjacent stable population from the Santa Cruz harbor (Reem et al., 2013a), and with other USA West-Coast populations as well. Thus, the genetic structures, possible genetic drifts, mutation rates and gene flow of the EYC populations are evaluated in this manuscript, following seven sampling sessions (between 1996 and 2008). Using 5 microsatellite markers as well as mitochondrial COI gene sequences, we also aim to define historic gene flow. Identification of past population dynamics may help understand the influences and risks of future species expansions, as of invasive potential.

2. Materials and methods

2.1. Location

The EYC area (36°48′49″N–121°47′13″W; Fig. 1) is connected to the Elkhorn Slough coastal estuary, an estuary drains into the Pacific Ocean through the Moss Landing Harbor Channel (http://www.montereybay.com/creagrus/elkhornslough.html). Once every few years, during wintertime, the EYC area undergoes major freshwater flooding catastrophes from one or more sources (Elkhorn Slough, Moro Cojo Slough, and/or Old Salinas River), which eradicate the marine biota in EYC. According to the Californian flood history reports (http://www.water.ca.gov/sfmp/resources/Attachment_C_History_Appendices_A-F.pdf; Fig. S1), the area experienced several major flooding catastrophes during 1996–2008 (Table S1).). One of the significant events throughout the studied period occurred during the winter of 2005 (http://met.nps.edu/~ldm/renard_wx/; the highest precipitation levels recorded for the period of 1996–2008), during which *B. schlosseri* colonies were not found in EYC (along with other marine organisms) due to the resulting prolonged period of low salinity (http://www.water.ca.gov/sfmp/resources/Attachment_C_History_Appendices_A-F.pdf).

2.2. Collection and preparation of genetic material

Colonies of *B. schlosseri* were collected from artificial objects (e.g., ropes, floats and buoys) and partly submerged floating docks, located 0.1–2.5 m below surface seawater. Six sampling sessions were performed, in 1996 ($n = 8$), 1999 ($n = 22$), 2000 ($n = 30$), 2001 ($n = 22$), 2007 ($n = 24$) and 2008 ($n = 20$) during the month of January, and an additional, single session was performed during November 1996 ($n = 15$), on *B. schlosseri* colonies situated at least 1 m apart, minimizing the chances of collecting kin colonies that may be identical by descent (Grosberg, 1987). In addition, colonial sampling was performed during January 2001 in five other USA West Coast populations, along more than 1520 km, including two marinas in the Seattle area, WA; the Shilshola Bay Marina (SM, $n = 28$) and the Des Moins Marina (DMM, $n = 28$; >1180 km north to EYC). Three sites were in California; the Santa Barbara Marina (SBM, $n = 26$; 334 km south of EYC), the Half Moon Bay Marina (HMBM, $n = 20$; 99 km north of EYC) and the Monterey Marina (MM, $n = 22$; 25 km south of EYC; Fig. 1). In total, 265 whole colonies or colonial fragments were removed from their substrates, and DNA extractions were carried out according to Paz et al. (2003). The vials were shipped to the National Institute of Oceanography, Israel, for further analysis.

2.3. Genotyping and sequencing

Microsatellites – Five *B. schlosseri* microsatellites, PBC-1, PB-29, PB-41, PB-49 (Stoner et al., 1997) and BS-811 (Pancer et al., 1994), were amplified using a polymerase chain reaction (PCR) with specific fluorescent primers (Agentech), following Reem et al. (2013a). Fragments were analyzed using an ABI-PRISM 310 sequencer with a 0.3 μl LIZ 500 size standard (Applied Biosystems, Foster City, CA) per well. Genotyping was carried out with the Genotyper software (Applied Biosystems, Foster City, CA).

mtDNA – The *B. schlosseri* cytochrome oxidase I (COI) gene (~700 bp) was amplified following Folmer et al. (1994; HCO2198-R, LCO1490-F, 5) in all the collected colonies (12 sampling sessions, 265 DNA samples). Sequencing was performed by Macrogen Inc. (Seoul, South Korea) on both forward and reverse directions.

2.4. Statistical analyses

Microsatellites – Analyses were performed on all yearly EYC samples and compared to the other West Coast USA samplings. Expected and observed heterozygosity (*He*, *Ho*), pairwise and overall genetic differentiation (θ, Weir and Cockerham, 1984) among EYC samplings/West Coast populations and their significance tests

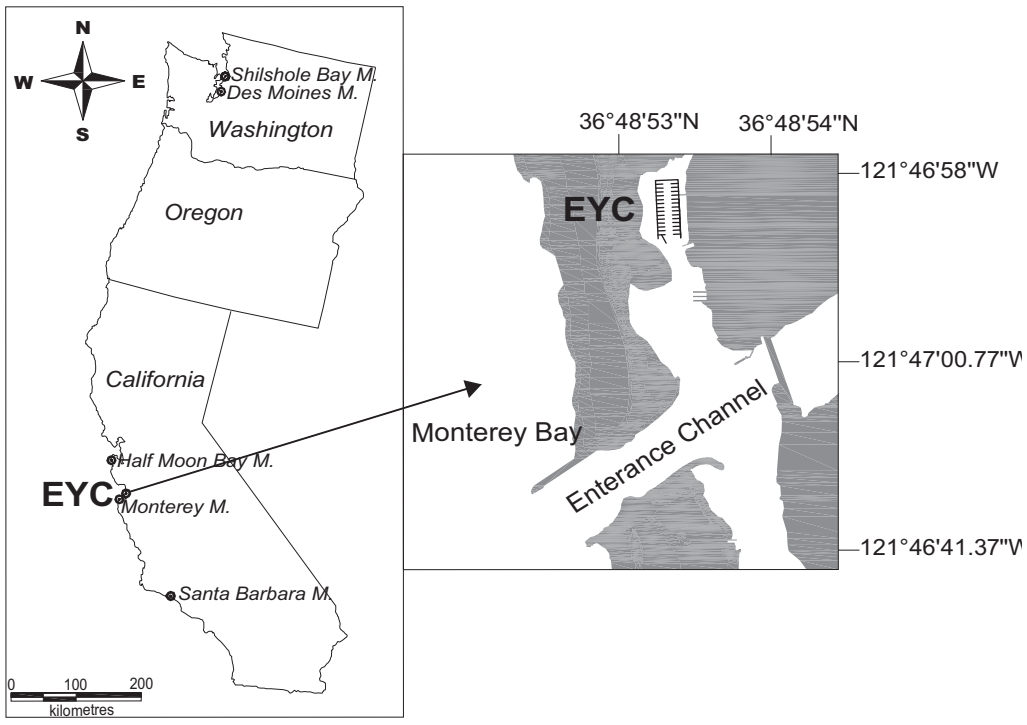

Fig. 1. *B. schlosseri* sampling locations along the USA West Coast; the Elkhorn Yacht Club (EYC-CA), the Shilshole Bay Marina (SM-WA), the Des Moines Marina (DMM-WA), the Santa Barbara Marina (SBM-CA), the Half Moon Bay Marina (HMBM-CA), the Monterey Marina (MM-CA). Right map: a/the detailed EYC location.

(10.000 permutation), linkage disequilibrium (Weir, 1979), population migration rates (pairwise *Nm* values) along with observed (*Ho*) and expected (*He*) (unbiased) heterozygosity and Nei's genetic distance (Nei, 1978) were calculated using GENETIX v.4.05 (Belkhir et al., 1996).

Pairwise and overall G_{ST} (a measure of genetic differentiation among subdivided populations for multiple alleles; Nei, 1973) and D_{est} (an estimator of actual differentiation; Jost, 2008) values were calculated by performing 9999 permutations for the EYC samplings and also for each population-locus combination; the number of different alleles (Na), of effective alleles (Ne) and of private alleles was calculated using GenAlEx 6.5 (Peakall and Smouse, 2012). Allelic richness (*AR*) and the inbreeding coefficient (F_{IS}; Weir and Cockerham, 1984) were estimated using FSTAT v.2.9.3 (Goudet, 1995). HP-Rare v.1.1 (Kalinowski, 2004, 2005) was used to calculate private allelic richness for each locus. The presence and frequency of null alleles was tested using the algorithm of Dempster et al. (1977), implemented in FreeNA (Chapuis and Estoup, 2007). The null hypothesis of the Hardy–Weinberg equilibrium at a locus was tested using the GENEPOP version 4.1.1 (Rousset, 2008). The analysis of the genetic relationships between the samples was estimated by Principle Component Analysis (PCA; Pearson, 1901) using PCA-GEN v.1.2. (Goudet, 1999).

The Bayesian structure analysis was run using STRUCTURE 2.1 (Pritchard et al., 2000). The number of populations assumed by STRUCTURE is *K*, each of which is in the Hardy–Weinberg and linkage equilibrium. Ten replicate simulations were run for each prior value of *K*. *K* was allowed to range from 1 to 10, using 30.000 burn-in repetitions and a final run of 10^6 Markov-chain-Monte-Carlo steps under an admixture model with independent allele frequencies. The most likely value for *K* was estimated with Evanno's Delta *K* method (Evanno et al., 2005) using the STRUCTURE HARVESTER (Earl and VonHoldt, 2012). CLUMPP software (Jakobsson and Rosenberg, 2007) was used to align the repetitions for each *K*. Mantel's (Mantel, 1967) statistic test was employed to understand the relationship between geographic distance and genetic divergence using the ARLEQUIN 3.11 (Excoffier et al., 2005).

Two time series samples data analysis: SCH (based on Reem et al., 2013a *Appendix Table 4) and EYC* – The Principle Component Analysis (PCA; Pearson, 1901) and the Agglomerative Hierarchical Clustering analysis (AHC; Ward, 1963) were performed using Euclidian distance by XLSTAT (Addinsoft™ XLSTAT v. 2013.6.01, Paris, France, 2014) in order to observe any dissimilarity between the SCH and EYC time series samples.

mtDNA – For COI gene analyses, DNA sequences were aligned in the BIOEDIT version 7.0.9.0 (Hall, 1999) software. Good quality 239 sequences were used for a/the NBLAST analysis (http://blast.ncbi.nlm.nih.gov/Blast.cgi) for identified organisms already found in the Gene bank, and further confirmed by performing a comparison to the BOLD (http://v2.boldsystems.org/views/login.php) database. Pairwise and overall genetic distances (θ, Weir and Cockerham, 1984) among EYC samplings/West Coast populations and their significance test were calculated by performing 10.000 permutations of the dataset using ARLEQUIN 3.11 (Excoffier et al., 2005).

Demographic history and neutrality tests – All analyses were performed using the DNAsp version 5.0 software (Rozas and Rozas, 1999; Rozas et al., 2003). The number of polymorphic sites (Np), the number of haplotypes (Nh), nucleotide diversity (Pi), haplotype diversity (Hd) and the gene flow parameter (*Nm*) between the time series EYC samples were estimated (Nei, 1973). The signatures of population demographic changes in all samplings/populations was investigated with a test of neutrality, Tajima's *D*-test (Tajima, 1989, using 1000 simulated samples), and compared with Fu and Li's D^* and F^* (Fu and Li, 1993) test (1000 simulations), further considered as a more sensitive metric (Fu, 1997). The distribution of pairwise nucleotide differences (mismatch distribution) of the EYC samples was calculated as an additional test for demographic expansion, using the Raggedness Index (*r*; Harpending, 1994).

COI based phylogeographic analyses – The phylogeographic relationships between haplotypes were estimated using the median joining algorithm with default settings for constructing the network (weight = 10 $\varepsilon = 0$) in the program NETWORK version 4.6.1.2. (Bandelt et al., 1999) and visualized with the Network Publisher (Fluxus Engineering, Clare, UK). The rho (ρ) estimator (Forster et al., 1996; Saillard et al., 2000) was used to date the haplotype cluster. The average number of mutations separating ancestral and descendent haplotypes was used to estimate the divergence time. The standard deviation sigma (SDs) and the standard deviation in a year (SDy) were also calculated. All were done using default NETWORK program parameters.

The Bayesian analysis was conducted using MrBayes v.3.21 (Ronquist and Huelsenbeck, 2003). Two independent runs were carried out with Monte-Carlo Markov Chains for 20 million generations, started from a random tree and sampled every 2.000 generations. The convergence of the analyses was validated by the standard deviation of split frequencies (<0.01) and by monitoring the likelihood values over time using TRACER v1.5 (http://beast.bio.ed.ac.uk/Tracer). The final models were visualized with FIGTREE v.1.3.1 (http://tree.bio.ed.ac.uk/software/figtree) and by default the initial 25% of all sampled trees were discarded.

3. Results

3.1. Microsatellites

The summarized statistics of 141 EYC (1996–2008, seven samplings) and 124 (six populations, year 2001) Californian/Seattle *B. schlosseri* colonies are given in Table 1. All five microsatellite loci were highly polymorphic. Locus BS-811 has emerged as the most polymorphic, with 55 alleles (consistent with results obtained from other *B. schlosseri* populations worldwide; Paz et al., 2003; Ben-Shlomo et al., 2006, 2010; Reem et al., 2013a, 2013b), while PBC-1, PB-29, PB-41 and PB-49 revealed in total 17, 8, 15 and 7 alleles, respectively, altogether presenting 113 scored alleles for all regions and samplings (Table 1, raw data is given in Appendix B).

No significant linkage genotypic disequilibrium was found between all pairs of the 5 microsatellite loci used (Bonferroni's correction, *P*-value = 0.01).

3.1.1. EYC time series samples (1996–2008)

3.1.1.1. Allelic structures. Allelic scores for seven samplings on loci PBC-1, PB-29, PB-41, PB-49 and BS-811 were 13, 5, 13, 13 and 36 respectively; minimum allelic richness was recorded in the 01/1996 sample, with 5, 4, 4, 5 and 3 alleles, respectively, and maximal allele numbers were obtained in the year 2007 with 8, 4, 7, 12 and 15 alleles, respectively. Locus BS-811 exhibited high F_{IS} and null allele ($r \geqslant 0.2$) values at all samplings/populations, except for EYC-01/1996. The high F_{IS} values were also observed at locus PB-41 at all EYC samplings (except 01/1996 and 11/1996) and at locus PB-49 during 2000 and 2007.

The most common microsatellite alleles (Fig. 2), while being prominent throughout the entire seven sampling session, fluctuated between the different sampling dates, particularly in PBC-1 and BS-811. The highest frequency of the most common allele at locus PBC-1 (202 bp) was 0.615 in 11/1996; 0.583 for allele 157 bp at locus PB-29 (year 2007); 0.571 for allele 175 bp at locus PB-41 during 01/1996 and 11/1996; 0.500 for allele 231 bp at locus PB-49 (01/1996) and 0.286 (274 bp) for BS-811 during 01/1996 (Fig. 2). The frequency of the most common allele in locus BS-811 (274 bp) was reduced gradually, from 0.286 to 0.021 between 01/1996 and 2007, and then it increased within the following 12 months, up to 0.111 in 2008.

The highest number of alleles at the EYC site ($n = 46$) was observed in the 2007 samplings, whereas the effective numbers of alleles increased over time, except for a small decrease in 2007 (Fig. 3a). In total, minimal private allelic richness was documented in the 01/1996 samplings (0.226), compared to all other samplings presenting high private allelic richness (Fig. 3a, Table S2). The minimum private allelic richness frequencies for the loci PBC-1, PB-29 and PB-49 were 0.0, 0.0 and 0.008, respectively in the 11/1996 samples, and a frequency of 0.0 was observed in the 2007 samples at the PB-29 locus also; frequencies of 0.001 and 0.034 were observed for loci PB-41 and BS-811, respectively, in the 01/1996 samples. Maximum private allele riches were observed in the 2000 samples for loci PBC-1 (0.167) and PB-41 (0.674); in the year 2001, for loci PB-29 (0.191) and BS-811 (1.373); in the year 2007, for locus PB-49 (0.602). The most distinctive decrease in private alleles was observed between the 2007 samples and the 2008 samples at locus PB-49 (from 0.602 to 0.063).

3.1.1.2. Genetic differentiation among sampling sessions and gene flow. While no significant pairwise θ were observed between the time series samples (Table 2), the pairwise G_{ST} and D_{est} values for the 2001–2007 samples were ascertained as significant (0.019 and 0.125, respectively $P = 0.001$; Table 3). On the other hand, overall θ, G_{ST} and D_{est} values were observed to be significant after the Bonferroni correction was performed for the EYC samples ($\theta = 0.022$, $G_{ST} = 0.021$ and $D_{est} = 0.077$, $p < 0.01$).

The first two axes of PCA explained 52.44% of the total variance, representing a large proportion of the genetic variance for seven EYC samples (Fig. 4a), with the first component (PC1) separating EYC samples 11/1996, 2007 and 2008 from the other three samples (01/1996, 1999 and 2001). The 2001 and 2007 samples were separated from all others by the second component (PC2). Even with only a 10 month period between sessions 01/1996 and 11/1996, the samples were separated by PC1 (28.84%).

All gene flow estimations (Nm) between consecutive samplings were insignificant ($Nm > 4$), insufficient to prevent genetic differentiation ($Nm > 1$, there is enough gene flow to negate the effects of genetic drift, $Nm > 4$, local populations belong to one randomly mating population (Wright, 1931)). The highest values of mean gene flow were found between samplings 1999 and 2000 (999999.0, $\theta \leqslant 0$; -0.006). The lowest value of mean gene flow (5.80) was detected between samplings 01/1996 and 11/1996 (the shortest sampling interval). The gene flow between the year 2000 and 2001 and between the year 2007 and 2008 was 10.90 and 12.70, respectively.

3.1.1.3. The structure analysis of seven EYC samplings. Using Evanno et al.'s (2005) approach, the most likely number of clusters was determined as 3 (Fig. 5a; Fig. S2a), and the outcomes of the Bayesian structure analysis performed for $K = 3$ are summarized in Table S4. The bar plots are shown in Fig. 5a for $K = 3$ from the CLUMPP software (which aligns multiple runs of structure), from 10 runs at each K. Cluster-1 was dominant in the years 01/96, 11/1996, 1999 and 2008 (41%, 46%, 41% and 47%, respectively), the 2001 samples consisted mainly of Cluster-2 (72%), and Cluster-3 had maximum value in the 2007 samples (53%). EYC-2000 revealed the most homogeneous samples [Cluster-1: 27% (red), Cluster-2 (yellow): 34%, Cluster-3 (blue): 39%, Fig. 5a, Fig. S2a, Table S4].

3.1.2. The West Coast 2001 populations

Allele numbers of the PBC-1, PB-29, PB-41, PB-49 and BS-811 microsatellites for the year 2001 at the West Coast sampling sites were: 9, 5, 6, 7, 13 (EYC); 6, 4, 3, 5, 18 (SM); 9, 5, 5, 6, 11 (DMM); 8, 3, 4, 6, 15 (SBM); 8, 5, 7, 11, 13 (HMB) and 5, 6, 7, 4, 13 (MM),

Table 1
The summary of the genetic diversity indices performed on *B. schlosseri* colonies sampled from the Moss Landing area (EYC), the Shilshole Marina (SM), the Des Moines Marina (DMM), the Santa Barbara Marina (SBM), the Half Moon Bay Marina (HMBM) and the Monterey Marina (MM). *N*, sample size; *k*, number of alleles; *AR*, allelic richness; *He*, expected heterozygosity; *Ho*, observed heterozygosity; F_{IS}, inbreeding coefficient; *Null* frequency of null alleles under the hypothesis of the Hardy–Weinberg equilibrium.

Parameter	Locus					Parameter	Locus				
	PBC-1	PB-29	PB-41	PB-49	BS-811		PBC-1	PB-29	PB-41	PB-49	BS-811
EYC-01/1996 (N = 8)						*SM-2001 (N = 28)*					
k	5	4	4	5	3	*k*	6	4	3	5	18
AR	5.000	3.992	4.000	5.000	3.000	*AR*	4.702	3.130	2.520	3.524	8.762
He	0.780	0.708	0.648	0.736	0.703	*He*	0.768	0.541	0.505	0.465	0.902
Ho	0.571	0.250	0.429	0.143	0.000	*Ho*	0.500	0.357	0.130	0.200	0.259
F_{IS}	0.284	0.663	0.357	0.818	1.000	F_{IS}	0.353	0.344	0.746[a]	0.574	0.717[a]
Null	0.065	0.261[b]	0.094	0.330[b]	0.396[b]	*Null*	0.145	0.109	0.254[b]	0.196	0.354[b]
EYC-11/1996 (N = 15)						*DMM-2001 (N = 28)*					
k	5	3	5	6	8	*k*	9	5	5	6	11
AR	4.218	2.759	3.886	4.978	6.803	*AR*	6.261	4.186	3.584	4.617	7.001
He	0.600	0.582	0.619	0.739	0.871	*He*	0.847	0.741	0.526	0.733	0.861
Ho	0.462	0.714	0.286	0.500	0.077	*Ho*	0.643	0.208	0.167	0.261	0.167
F_{IS}	0.238	−0.238	0.548	0.333	0.915[a]	F_{IS}	0.244	0.723[a]	0.688[a]	0.649[a]	0.810[a]
Null	0.050	0.000	0.205[b]	0.103	0.414[b]	*Null*	0.145	0.305[b]	0.250[b]	0.270[b]	0.367[b]
EYC-1999 (N = 22)						*SBM-2001 (N = 26)*					
k	9	4	6	9	8	*k*	8	3	4	6	15
AR	6.142	3.527	4.437	6.180	6.380	*AR*	5.026	2.844	3.197	4.528	7.524
He	0.802	0.693	0.735	0.783	0.852	*He*	0.691	0.499	0.613	0.618	0.866
Ho	0.591	0.818	0.333	0.550	0.211	*Ho*	0.308	0.539	0.462	0.423	0.539
F_{IS}	0.268	−0.185	0.553[a]	0.303	0.758[a]	F_{IS}	0.559[a]	−0.082	0.251	0.320	0.383[a]
Null	0.109	0.017	0.231[b]	0.120	0.339[b]	*Null*	0.229[b]	0.000	0.065	0.132	0.174
EYC-2000 (N = 30)						*HMBM-2001 (N = 20)*					
k	8	4	9	9	15	*k*	8	5	7	11	13
AR	5.404	3.800	6.039	6.023	8.729	*AR*	5.386	3.758	5.432	7.529	8.052
He	0.715	0.721	0.758	0.798	0.913	*He*	0.772	0.591	0.818	0.881	0.895
Ho	0.679	0.552	0.250	0.500	0.179	*Ho*	0.800	0.400	0.550	0.579	0.450
F_{IS}	0.052	0.237	0.675[a]	0.378[a]	0.807[a]	F_{IS}	−0.04	0.329	0.333	0.349	0.504[a]
Null	0.057	0.120	0.323[b]	0.160	0.381[b]	*Null*	0.021	0.156	0.139	0.167	0.228[b]
EYC-2001 (N = 22)						*MM-2001 (N = 22)*					
k	9	5	6	7	13	*k*	5	6	7	4	13
AR	5.899	3.942	4.763	5.237	8.474	*AR*	3.609	4.577	5.427	3.918	7.742
He	0.799	0.637	0.717	0.760	0.912	*He*	0.635	0.711	0.769	0.744	0.865
Ho	0.773	0.364	0.364	0.682	0.524	*Ho*	0.810	0.571	0.476	0.450	0.318
F_{IS}	0.034	0.435	0.499[a]	0.105	0.432[a]	F_{IS}	−0.283	0.200	0.387	0.401	0.637[a]
Null	0.020	0.222[b]	0.208[b]	0.050	0.195 [b]	*Null*	0.00	0.062	0.171	0.173	0.287[b]
EYC-2007 (N = 24)											
k	8	4	7	12	15						
AR	4.670	3.177	5.134	7.645	7.594						
He	0.661	0.582	0.773	0.882	0.844						
Ho	0.583	0.583	0.375	0.550	0.250						
F_{IS}	0.120	−0.002	0.520[a]	0.383[a]	0.708[a]						
Null	0.003	0.000	0.221[b]	0.192[b]	0.317 [b]						
EYC-2008 (N = 20)											
k	7	4	5	7	12						
AR	4.66	3.59	4.65	5.36	8.65						
He	0.612	0.671	0.770	0.788	0.922						
Ho	0.316	0.500	0.353	0.563	0.278						
F_{IS}	0.491[a]	0.261	0.549[a]	0.293	0.705[a]						
Null	0.192	0.127	0.229[b]	0.106	0.330[b]						
Overall EYC						*Overall all sites*					
k	13	5	13	13	36	*k*	17	8	15	18	55
AR	5.64	3.60	5.43	6.15	9.45	*AR*	6.10	3.80	5.30	6.20	10.03

[a] Significant F_{IS} value after Bonferroni correction for multiple tests (Rice, 1989).
[b] High null allele value ($r \geqslant 0.2$). PBC-1, PB-29, PB-41, PB-49 from Stoner et al. (1997) and BS-811 from Pancer et al. (1994).

respectively. For locus BS-811 we observed high F_{IS} and null allele values ($r \geqslant 0.2$) at all the 2001 samples except the SBM-2001 samples. Similar high F_{IS} values were observed in locus PBC-1 at SBM, locus PB-29 at DMM, locus PB-41 at SM and DMM, and locus PB-49 at DMM (Table 1). The highest private allelic richness was observed at locus BS-811 for the SM, DMM, SBM, HMBM and MM samples (1.654, 0.928, 0.454, 1.296 and 1.174, respectively; Fig. 3b, Table S2). Significant θ values ($p < 0.0017$) were also observed between all the sampled 2001 populations, ranging from 0.036 (between EYC-01 and HMBM) to 0.227 (between SM and SBM; Table S3).

The first two PCA axes explained 72.52% of the total variance for the six 2001 West Coast populations (Fig. 4b). The first component (PC1 = 51.24%) clearly outlined the separation of the two Washingtonian populations (SM, DMM) from the four Californian populations (SBM, HMBM, MM, EYC-01). PC2 (21.28%) revealed that the DMM and HMBM samples are separated from the other populations.

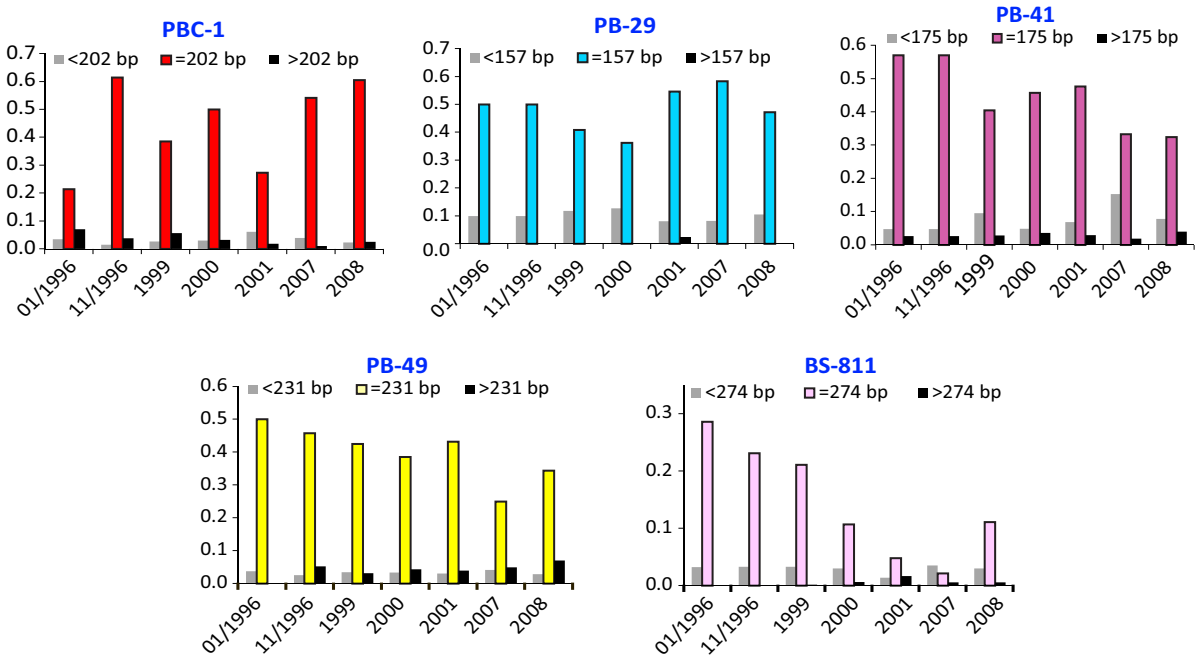

Fig. 2. Allelic frequencies (from five microsatellite loci) of the time series collections (1996–2008) at the Elkhorn Yacht Club (EYC). Legends: allele ID = allele size (bp) shows the frequency of a specific microsatellite locus most common allele; allele ID < allele size (bp) depicts the mean value of all the alleles of a specific microsatellite locus which have a smaller "bp" value than the most common allele; allele ID > allele size (bp) depicts the mean value all alleles of a specific microsatellite locus which have a bigger "bp" value than the most common allele. Locus identification is marked above each sub-fig.

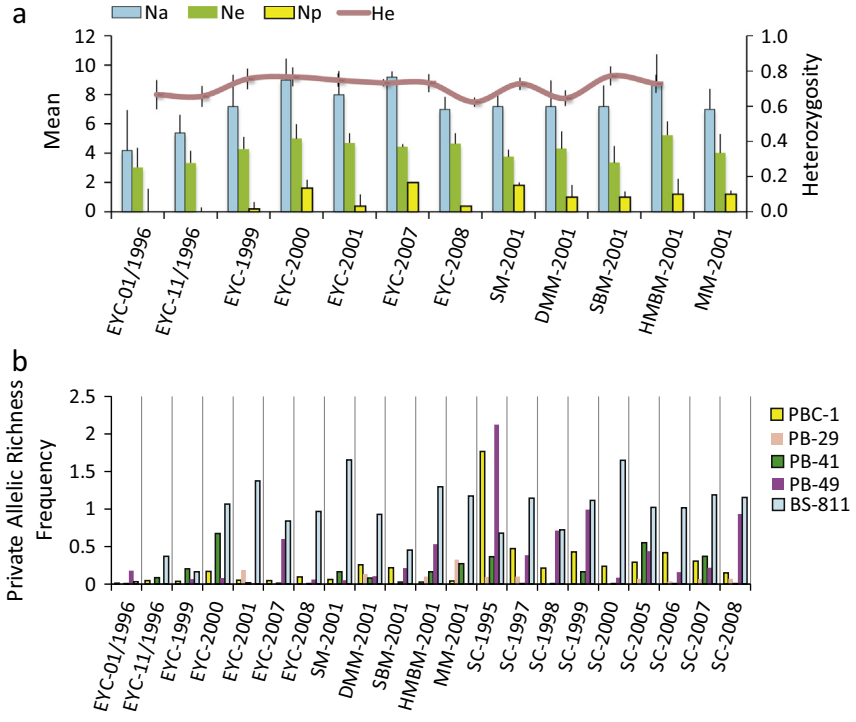

Fig. 3. Population genetic characteristics for *B. schlosseri* populations along the USA Pacific coasts. (a) Effective alleles, expected heterozygosity and private alleles for *B. schlosseri* populations sampled in the present work, from 6 sites (the EYC-time series, SM-2001, DMD-2001, SBM-2001, HMBM-201, and MM-2001). (b) Private allelic richness frequencies for 6 sites (WA and CA) and SCH (27 km south of EYC; SCH data taken from Reem et al., 2013a). Na: the number of different alleles, Ne: the number of effective alleles, Np: the number of private alleles, *He*: the expected heterozygosity; see Table S2 for details.

The outcomes of the Bayesian structure analysis for the most likely number of clusters ($K = 2$ and $K = 5$), which are characterized with high Delta K and a low SD value, are summarized in Table S4. According to $K = 2$, Cluster-1 consisted mainly of the northern populations (SM and DMM; 92% and 72%, respectively), and Cluster-2 of the southern populations (SBM, HMBM, MM and EYC; 94%, 80%, 78% and 87%, respectively). According to $K = 5$, the SM samples had maximum values in cluster-5 (71%), the DMM at

Table 2
Pairwise genetic differentiation (θ) estimates between the EYC time series samples for 5 microsatellites loci (θ-Mic: above diagonal) and the mitochondrial COI gene (θ-COI; below diagonal). Pairwise θ were calculated using the method of Weir and Cockerham (1984), and were tested using 10.000 permutations. All values became non-significant after a sequential Bonferroni correction.

θ-Mic/θ-COI	01/1996	11/1996	1999	2000	2001	2007	2008
01/1996	0	0.042^{NS}	-0.002^{NS}	0.004^{NS}	0.035^{NS}	0.057^{NS}	0.045^{NS}
11/1996	0.184^{NS}	0	0.025^{NS}	0.019^{NS}	0.040^{NS}	0.029^{NS}	0.007^{NS}
1999	0.027^{NS}	0.015^{NS}	0	-0.006^{NS}	0.022^{NS}	0.033^{NS}	0.013^{NS}
2000	0.057^{NS}	-0.018^{NS}	-0.025^{NS}	0	0.023^{NS}	0.021^{NS}	0.012^{NS}
2001	0.236^{NS}	-0.018^{NS}	0.043^{NS}	0.011^{NS}	0	0.038^{NS}	0.033^{NS}
2007	0.118^{NS}	-0.031^{NS}	-0.024^{NS}	-0.025^{NS}	0.004^{NS}	0	0.019^{NS}
2008	0.222^{NS}	-0.025^{NS}	0.035^{NS}	0.003^{NS}	-0.045^{NS}	-0.004^{NS}	0

Bonferroni correction (Rice, 1989) have been done for multiple tests microsatellite and COI data ($P < 0.0012$), EYC; Elkhorn Yacht Club.

Table 3
Pairwise G_{ST} (above the diagonal) and D_{est} (below the diagonal) values based on 5 microsatellites loci for the EYC sampling sessions, calculated by performing 9999 permutations. Pairwise value between the 2001 and 2007 populations became significant after a Bonferroni correction.

G_{ST}/D_{est}	01/1996	11/1996	1999	2000	2001	2007	2008
01/1996	0	0.020^{NS}	-0.002^{NS}	0.002^{NS}	0.016^{NS}	0.029^{NS}	0.023^{NS}
11/1996	0.106^{NS}	0	0.013^{NS}	0.010^{NS}	0.021^{NS}	0.015^{NS}	0.004^{NS}
1999	-0.012^{NS}	0.075^{NS}	0	-0.003^{NS}	0.011^{NS}	0.017^{NS}	0.007^{NS}
2000	0.012^{NS}	0.060^{NS}	-0.018^{NS}	0	0.012^{NS}	0.011^{NS}	0.006^{NS}
2001	0.103^{NS}	0.116^{NS}	0.078^{NS}	0.082^{NS}	0	0.019^{*}	0.017^{NS}
2007	0.177^{NS}	0.081^{NS}	0.113^{NS}	0.075^{NS}	0.125^{*}	0	0.010^{NS}
2008	0.143^{NS}	0.021^{NS}	0.046^{NS}	0.041^{NS}	0.110^{NS}	0.062^{NS}	0

Permutation 9999. Significant values at $P < 0.0012$ (Bonferroni correction for multiple tests microsatellite, Rice, 1989).

Cluster-4 (57%), the SBM at Cluster-2 (56%), the HMBM at Cluster-1 (47%) and the MM samples at Cluster-3 (54%). The EYC-2001 samples had the most homogeneous pattern (Cluster-1: 20%, Cluster-2: 34%, Cluster-3: 36%, Cluster-4: 5% and Cluster-5: 5%, Fig. 5b, Fig. S2b, Table S4).

The Mantel test detected a significant positive relationship between pairwise genetic and geographical distances for the entire region (mean value $\theta = 0.07$, sums of squares $\theta = 0.2$, $P = 0.001$).

3.1.3. The SCH time series samples

The nearest populations to EYC are at SCH, a site studied intensively for its *B. schlosseri* population genetics during 1995–2008 (Reem et al., 2013a), overlapping the current research period (1996–2008). We obtained the genotype data for loci PBC-1, PB-29, PB-41, PB-49 and BS-811 from the study of Reem et al. (2013a). Loci PBC-1 and BS-811 were discarded because we were unable to establish cross-study relationships between alleles at this locus. Alleles at the other three loci were pooled (Table S5). The resulting matrix of allele frequencies per sample (Table S5) was subjected to the PCA. The outcomes of the PCA on the allele frequency data are presented in Fig. 4c. It is thus a fact of significant interest that these closely situated populations differed significantly, as shown by the PCA (Fig. 4c; a 59.06% dissimilarity was seen by the first two axes) and the AHC analysis (Fig. S3, Table S5). Three clusters were revealed, one consisting of only the SCH-1995 samples, the second of the EYC-01/1996, -11/1996, -1999, -2000, -2001, SCH-1998 and -2006 and the third of the SCH-1997, -1999, -2000, -2005, -2007, -2008, EYC-2007 and -2008 samples (Fig. 4c).

3.2. Mitochondrial DNA analyses

The final length of the COI gene fragment after alignment and trimming was 594 bp. The main parameters describing populations, namely the number of polymorphic sites (Np), haplotypes (Nh), haplotype diversity (Hd) and nucleotide diversity (Pi), are summarized in Table 4. Altogether, 23 haplotypes were revealed from the 239 sequences. Polymorphic sites, haplotype diversity and nucleotide diversity were calculated as 58, 0.663 and 0.012, respectively. Almost 83% of the haplotypes are unique.

3.2.1. The EYC time series samples

3.2.1.1. Genetic differentiation among samplings and gene flow. The total number of polymorphic sites for the 131 EYC colonial samples were 49, representing 13 haplotypes, with a 0.0048-nucleotide diversity and a haplotype diversity of 0.467. The lowest haplotype numbers were obtained in the 11/1996 and 2008 samples ($n = 2$), and 6 haplotypes was the record number in the year 2000 sampling.

Except samplings 01/1996 vs. 11/1996, the pairwise gene flow estimates for all other were insignificant ($Nm > 4$; for 1999 vs. 2000; 2000 vs. 2001 and 2007 vs. 2008 [consecutive years sampling pairs] were 49.33, 18.35 and 21.23, respectively). The Nm value for samplings 01/1996 vs. 11/1996 was 2.95, which negates the effects of genetic drift but infers non-random mating ($Nm > 1$). During January 1996 only eight *B. schlosseri* colonies were sampled in the EYC area, possibly due to the 1995 January and March rainfalls, where the Salinas River exceeded its previous measured record crest by more than four feet, which was within a foot or two of the reputed crest of the legendary 1862 flood (http://www.water.ca.gov). When compared with the other consecutive samplings, the documented lowest gene flow between 01/1996 and 11/1996 could be explained as a result of these rainfalls and the following fresh-water floods.

3.2.1.2. Demographic history and the neutrality test. Based on Tajima's *D* and Fu's *Fs* tests, the null hypothesis for the COI gene neutral evolution was not rejected for most samplings, except for the EYC sessions for the years 2000 and 2007, which showed significant negative values, indicating population expansion, natural selection, demographic and/or geographic expansions (e.g., following the natural catastrophe that occurred during the winter of 2005 and the high precipitation of 2000, Fig. S1), (Tajima's *D*: −2.27, −2.50 and Fu and Li's *F*: −2.85, 4.15; Table 4). The total EYC COI data revealed the possibility of selective sweep (Tajima's *D*: −2.16, $P < 0.02$), population expansion or background selection (Fu and Li's *D* test statistic: −3.48, $P < 0.02$). Fu and Li's *F* test

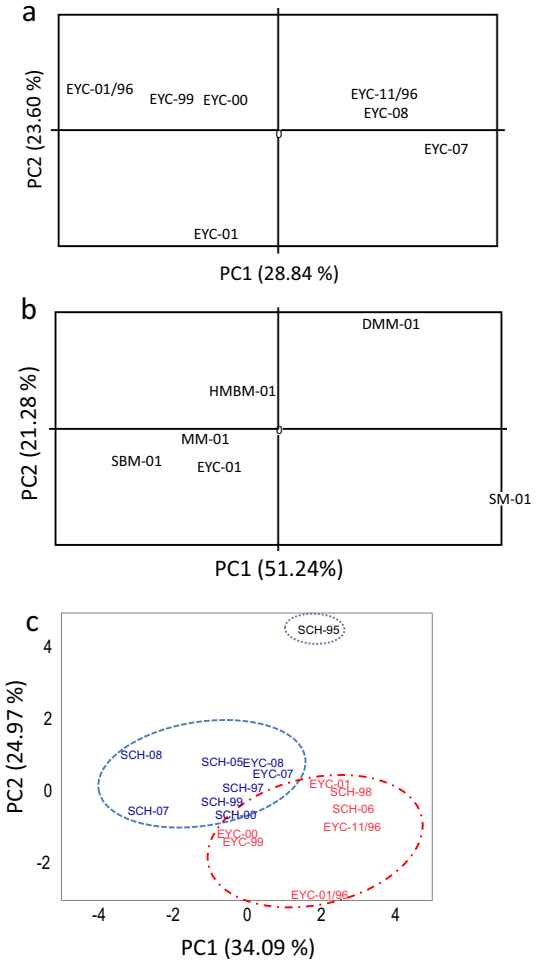

Fig. 4. Projection on the principal plane of PCA (PC1 and PC2), defined by their allelic frequencies at 5 microsatellite loci for: (a) The time series EYC samples: -01/1996, -11/1996, -1999, -2000, -2001, -2007, -2008 and (b) six 2001 *B. schlosseri* populations: SM, DMM, SBM, HMBM, MM and EYC. (c) The time series EYC (-01/1996, -11/1996, -1999, -2000, -2001, -2007, -2008) and SCH samples (-1995, -1997, -1998, -1999, -2000, -2005, -2006, -2007 and -2008; from Reem et al. (2013a)). Dashed lines were made according to following the hierarchical cluster analysis.

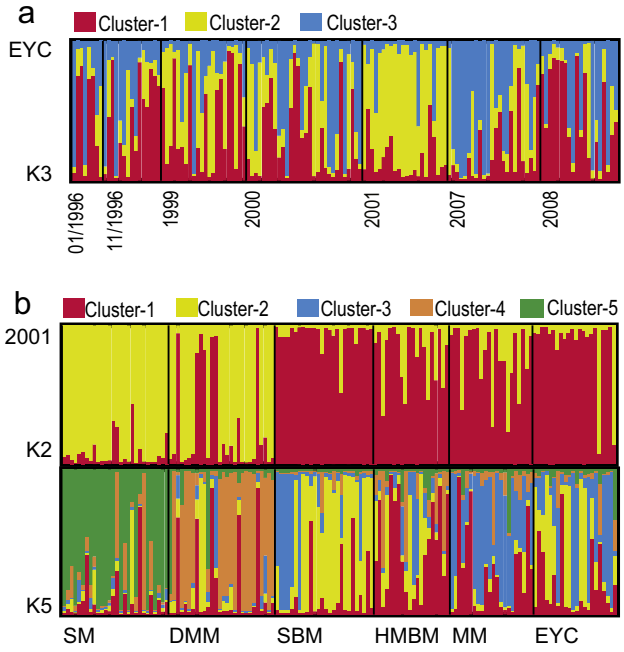

Fig. 5. The Bayesian structure analysis results for individual genotypes in the EYC and 2001 *B. schlosseri* samplings/populations from the USA Western Pacific coastline, based on 5 microsatellite loci. Bar plots from CLUMPP results aligned from 10 structure runs for (a) the EYC time series samples ($K = 3$) and (b) the 2001 samples ($K = 2$ and $K = 5$). Note: A single vertical line represents each individual with an estimated membership in each cluster, denoted by the different colors (see Table S4 for details of membership coefficients). Individuals are separated based on their sampling/population and the black vertical lines in the bar chart are sampling/population identifiers. Populations are ordered as per their population name. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

statistics (−3.51, $P < 0.02$) further indicated genetic hitchhiking in addition to population expansion. These tests showed that selective sweep, background selection and/or genetic hitchhiking can be inferred from the mtDNA patterns of the EYC samples (Table 4).

All distributions displayed a non-significant raggedness index ($r > 0.03$). These results were confirmed by the observed mismatch distributions. Unimodal and multimodal distributions were observed, fluctuating between 0.115 (01/1996) and 0.270 (11/1996; Fig. 6a and b), that is, occurring during the shortest time period (10 months) between two consecutive samplings. The mismatch distribution plot for the EYC-11/1996, 2001 and 2008 samples was smooth and unimodal (Fig. 6b,e and g), indicating a population expansion, whereas the multimodal pattern of the EYC-01/1996, 1999, 2000 and 2007 samples' mismatch distributions may suggest population subdivision and a stable population size (Fig. 6a,c,d and f).

3.2.2. The West Coast 2001 samples

When the neutrality test was performed for all 2001 *B. schlosseri* populations except the SM and HMBM samples, no selective sweep was revealed. The *D* and F^* tests of the SM and HMBM samples also indicated the possibility of population expansion, background selection or genetic hitchhiking (Table 4). Minimum polymorphic sites were observed at MM (Np = 0), and maximum in the SBM samples (Np = 33). The population structuring θ values (Bonferroni correction $P < 0.0017$) were significant for the USA West Coast samples, except for the pairs SBM-HMBM, DMM-SBM, DMM-HMBM and MM-EYC-2001 (Table S3, below diagonal).

3.2.3. Phylogeographic relationship among haplotypes

Relationships and the geographical distribution of the 23 haplotypes (239 West Coast *B. schlosseri* samplings) revealed a single mtDNA clade (named clade α and containing 14 haplotypes), separated by 11 mutations from the closest haplotype (H6) out of the other 9 haplotypes (Figs. 7a and b and 8), and containing 6 unique EYC haplotypes out of the 12 total unique haplotypes. The two most abundant haplotypes (H1, H4) were found in all 6 USA West Coast *B. schlosseri* populations. Except for the MM site, all other *B. schlosseri* samples were composed primarily of Haplotype-1, where Haplotypes-4 was primarily featured in the EYC samples and also in the HMBM and MM samples. Two semi-private haplotypes (H-2 and H-5), each containing 2–3 populations, were the most abundant haplotypes of clade-α. Haplotype-2 dominated the SM samples (89.47%) and was also found in the SBM and DMM populations. Haplotype-5 occurred primarily in the SBM and HMBM samples (Fig. 7b). According to network age estimation, clade-α ($\rho = 8.5625$) diverged from the other haplotypes approximately 240,842 years ago (SDs = 3.3, SDy = 65,697); H2 and H5 (ρ: 8.5625) diverged ca. 172,791 years ago (SDs = 1.1, SDy = 22,314); H1 and H4 ($\rho = 1.3495$) diverged 27,232 years ago (SDs = 1.1, SDy = 21,377). The Bayesian analysis revealed H1, H20 and H22 as the deduced source for the other COI haplotypes.

Table 4
COI gene diversity parameters calculated for the 12 *B. schlosseri* USA West Coast sites/dates.

Populations	*n*	Np	Nh	Pi	Hd	F&LD*	F&LF*	Taj. *D*
EYC-01/1996	10	21	5	0.0135	0.800	0.99^{NS}	0.94^{NS}	0.37^{NS}
EYC-11/1996	13	1	2	0.0005	0.282	0.73^{NS}	0.54^{NS}	-0.27^{NS}
EYC-1999	19	24	4	0.0078	0.509	0.94^{NS}	0.35^{NS}	-1.26^{NS}
EYC-2000	27	40	6	0.0069	0.496	-2.49^{NS}	-2.85^{a}	-2.27^{b}
EYC-2001	21	3	3	0.0010	0.495	-1.28^{NS}	-1.28^{NS}	-0.68^{NS}
EYC-2007	23	27	4	0.0044	0.383	-4.01^{b}	-4.15^{b}	-2.50^{c}
EYC-2008	18	1	2	0.0007	0.425	0.67^{NS}	0.82^{NS}	0.87^{NS}
Total EYC	131	49	13	0.0048	0.467	-3.48^{b}	-3.51^{b}	-2.16^{b}
SM-2001	27	26	4	0.0197	0.533	1.44^{a}	2.14^{b}	2.66^{b}
DMM-2001	19	27	3	0.0081	0.205	0.43^{NS}	-0.15^{NS}	-1.49^{NS}
SBM-2001	22	33	7	0.0192	0.671	0.62^{NS}	0.86^{NS}	1.00^{NS}
HMBM-2001	20	19	5	0.0159	0.695	1.29^{NS}	2.05^{b}	2.89^{b}
MM-2001	20	0	1	0.0000	0.000	–	–	–
Total all	239	58	23	0.0124	0.6628	-4.22^{b}	-3.19^{b}	-0.76^{NS}

n, sample size; Np, number of segregation sites (polymorphic sites); Nh, number of haplotypes; Pi, nucleotide diversity; Hd, haplotype diversity. D^* test statistic is based on the differences between the number of mutations appearing only once among the sequences (singletons), and the total number of mutations (Fu and Li, 1993). The F^* test statistic is based on the differences between the number of singletons and the average number of nucleotide differences between pairs of sequences (Fu and Li, 1993). Window length: 100 Step size: 25, $P < 0.10^{a}$, $P < 0.05^{b}$, $P < 0.02^{c}$, ^{NS}: not significant). EYC; Elkhorn Yacht Club, SB: Shilshole Bay, DMM; Des Moines Marina, SBM; Santa Barbara Marina, HMBM; Half Moon Bay Marina, MM; Monterey Marina.

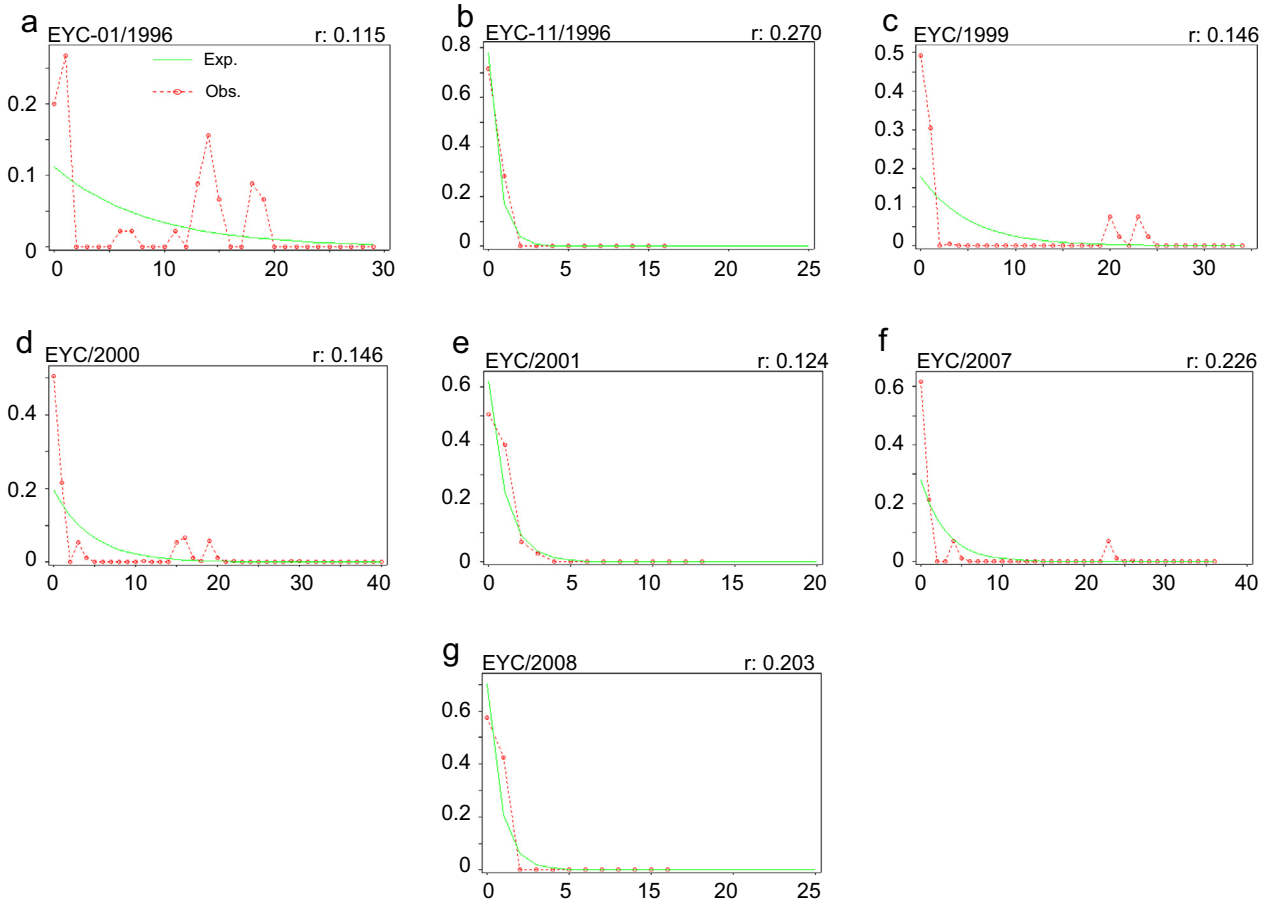

Fig. 6. Mismatch distribution based on the COI gene for the seven (a–g) EYC *B. schlosseri* samples. The *x*-axis represents the number of uncorrected pairwise differences and the *y*-axis represents frequency. Exp: Expected value, Obs: Observed value for constant population size.

Though H4 is close to H1, H2 and H5 are far removed from them and are included in clade-α (Fig. 8a). The posterior probabilities of the Bayesian analysis, which was conducted using MrBayes, are given in Fig. 8b.

4. Discussion

Botryllus schlosseri, while commonly used as a model species for studying allorecognition (Scofield et al., 1982), developmental

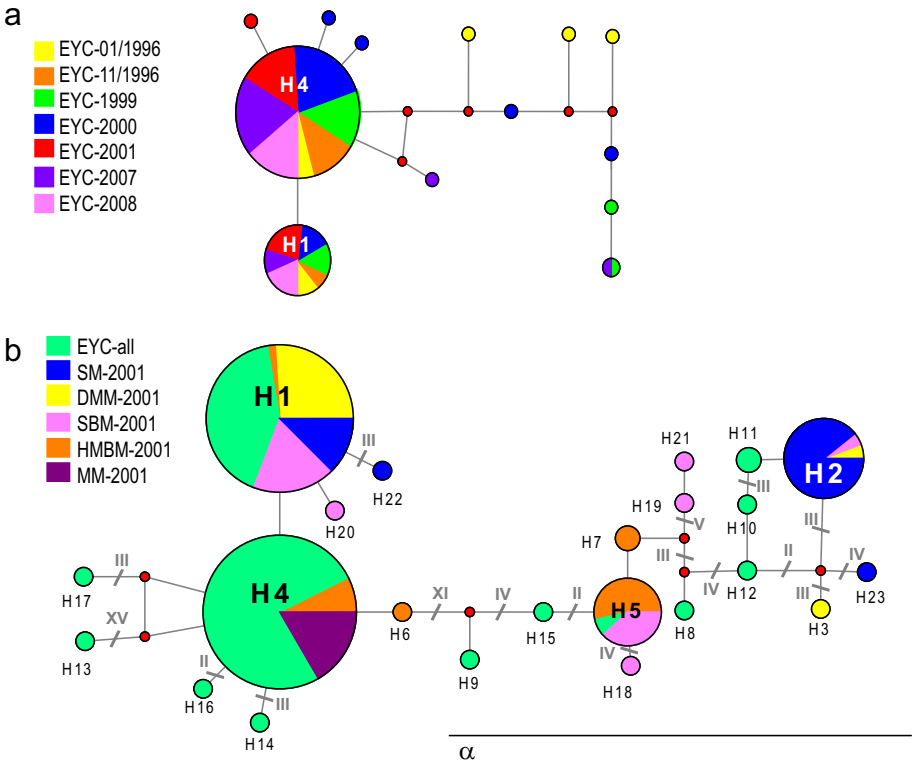

Fig. 7. The median-joining network of *B. schlosseri* for the COI haplotypes (H with a corresponding number denotes a specific haplotype). The pie size is proportional to the number of colonies, and colors indicate different samplings/populations/haplotypes. (a) The time series sampling for the 13 EYC haplotypes, revealing one mutation between the most common haplotypes, H1 and H4. (b) The network analysis of all 23 USA West Coast *B. schlosseri* haplotypes (all the EYC samples are clustered together, in green). Lines without Roman numerals: one mutation step between haplotypes; Roman numerals with vertical black lines: the number of mutations (>1). mv = median vectors; vertical black line = highlighted mutation positions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

biology (Rinkevich et al., 1992, 2007; Manni and Burighel, 2006), apoptosis (Tiozzo et al., 2006) and stem cell biology (Rosner et al., 2009; Rinkevich et al., 2013), is also a worldwide invasive species, known in the central Californian coastal waters for almost seven decades (Kozloff, 1975). Having a characteristic life history of five generations/year (Reem et al., 2013a) makes *B. schlosseri* extremely suited for population genetics research in California (Stoner et al., 2002; Bock et al., 2012; Reem et al., 2013a), also existence of several previously unrecognized and probably reproductively isolated cryptic *B. schlosseri* species/clades is possible (Bock et al., 2012).

Heterozygote deficiencies in *B. schlosseri* populations were recorded in this study for most sampling seasons and localities, resembling the results of former USA West Coast studies (Stoner et al., 2002; Ben-Shlomo et al., 2008; Reem et al., 2013a) as well as of *B. schlosseri* studies conducted elsewhere (Paz et al., 2003; Ben-Shlomo et al., 2006, 2010; Reem et al., 2013b). High F_{IS} can be explained by either asexual reproduction (irrelevant here due to the sampling protocol) or non-random mating between genetically related conspecifics. The results further documented high numbers of alleles per each microsatellite locus, rapid fluctuations in allelic frequencies between different years, high numbers of unique alleles and changes in the frequency of the most common alleles. The long term studies of Paz et al. (2003) and Reem et al. (2013a) have also revealed high mutation rates in the microsatellite alleles, further supported by the high fluctuations of the private allele frequencies, recorded here. In the Israeli Mediterranean populations, Paz et al. (2003) noted changes in microsatellite allele frequencies over a period of two years, including the appearance and the loss of alleles, probably caused by random genetic drift, a high mutation rate and high genetic variation that are all explained by kin mating and natural selection (Slatkin, 1995; Stoner et al., 2002). This high genetic diversity (also characteristic to other gene loci in botryllid ascidians; Yund and Feldgarden, 1992) is an attribute that may play a key role in the success of an invasive species following introductions into new habitats (Novak, 2007; Roman and Darling, 2007; Reem et al., 2013a). Studies have indeed pointed to positive relationships between microsatellite heterozygosity, fitness and genome-wide heterozygosity (Harrison et al., 2011), though such genetic scenarios have not yet been labelled for *B. schlosseri*.

These aforementioned results can be compared with Caputi et al.'s (2014) long term study (2002–2012) on the demographic and reproductive trends of *Ciona intestinalis* sp. A. Whereas precipitation appears to play no role in shaping the biological parameters of *Ciona intestinalis* sp. A, in the present study we observed a negative correlation between high precipitation and the population size of *B. schlosseri*, a result associated with the deeper living depth of the *Ciona intestinalis* populations (down to 500 m), whereas *B. schlosseri* is a relatively shallow water species (usually 0–60 m depth; unpubl.). Due to high precipitation in the catchment basins, several major floods in the Moss Landing area occurred during the 12 year period of this study, dramatically affecting marine biota in local marinas. The most pronounced effect was in the year 2005, where no single *B. schlosseri* sample could be found in EYC (the next sampling date was in 2007) and the entire community was replaced with fresh water organisms (BR personal observations). Theoretically, newcomers could have originated from neighboring population/s (SCH, HBMM, MM) and/or far away US West Coast populations. Another option, while less conceivable, is that deep-water *B. schlosseri* genotypes, located on/under the Moss Landing Harbor/Old Salinas River entrance rocks, may have escaped the 2005 low salinity event. Focusing on the PCA and HCA of the EYC and SCH time series samples and on the SCH-2007 and 2008

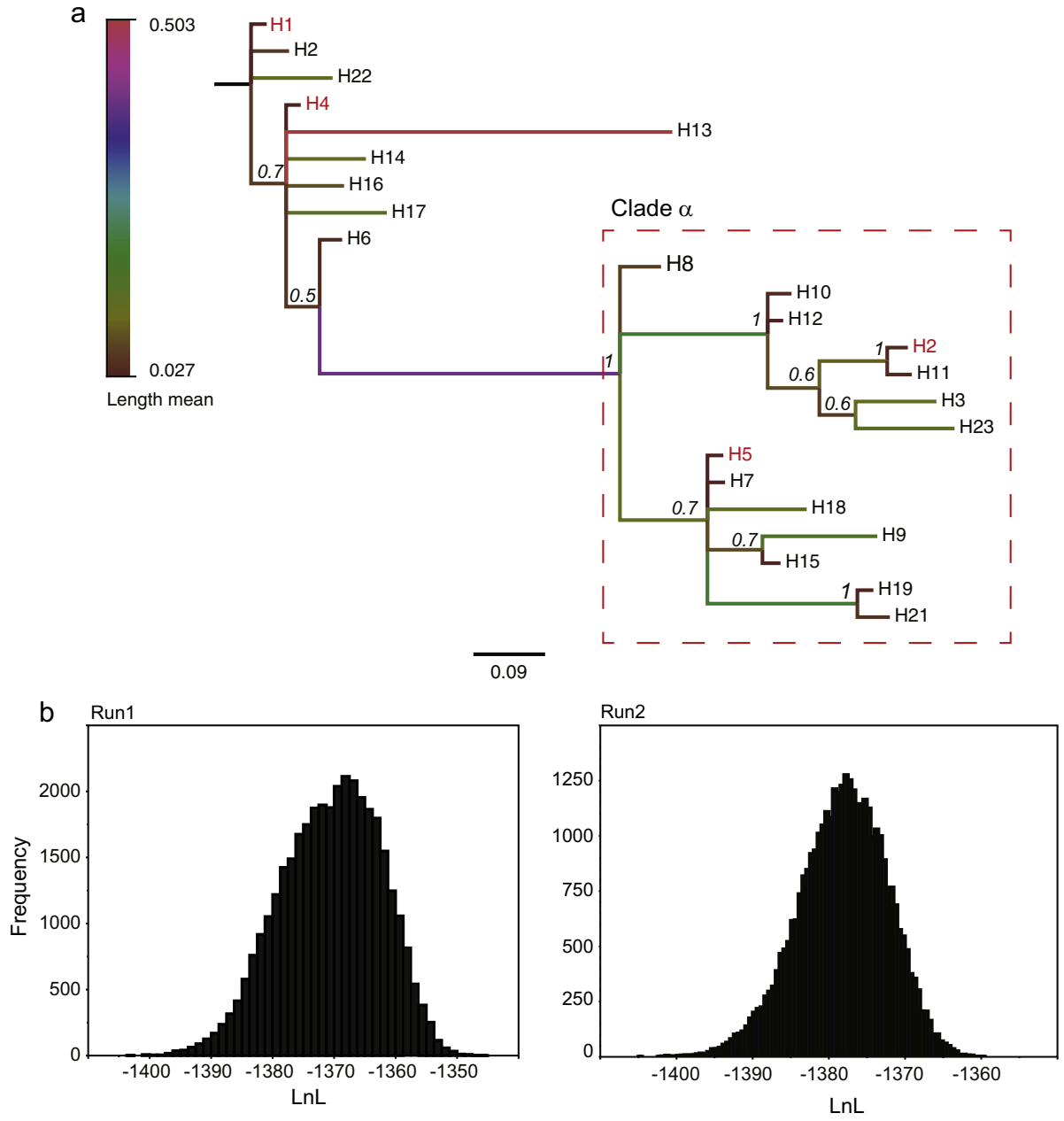

Fig. 8. (a) The Bayesian topology of the COI gene from a combined data set analysis shows the relationships between haplotypes of *B. schlosseri*. The branches' coloration is length encoded (brown for the shortest branch and red for the longest). Four major haplotypes are colored in red, values above nodes represent posterior probabilities. (b) The posterior probabilities of the Bayesian analysis for two runs from the Tracer v1.5 (http://beast.bio.ed.ac.uk/Tracer) (Summary Statistic: mean = −1370, −1378; std error of mean = 1.3, 3.7; 95% HPD lower = −1386, 1389; 95% HPD upper = −1355, 1367; auto-correlation time (ACT) = 5–727; effective sample size (ESS) = 38–24,737). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

samples (Reem et al., 2013a, Fig. 4c) made clear that the post-flood Elkhorn Yacht Club population was most probably established by the Santa Cruz Harbor migratory genotypes (hitchhiking on pleasure boats), and was further shaped by genetic drift and mutations. However, the contributions of populations from other marinas and of deep-sea survivors cannot be completely ignored.

Also relevant is the interesting fact that despite there being only 10 months between the EYC-01/1996 and 11/1996 samples, significant differences were recorded in the nuclear and mtDNA data of this site. This is a possible outcome of two smaller flood events that occurred during January–March 1995 and December–January 1996–1997 (Table S1), potentially altering the population genetics characteristic within a short time scale.

4.1. The *B. schlosseri* populations of the Pacific coasts

We recorded high θ values (ranging between 0.071 and 0.227 for microsatellites, 0.147 and 0.605 for COI) between Washingtonian (SM, DMM) and Californian (SBM, HMBM, MM, EYC) USA West Coast populations, hallmarked as well by the outlined separation of the Washingtonian populations from the Central/Southern Californian populations. It should be noted that the *B. schlosseri* populations in the Seattle area are relatively new (first recorded in the area in the early 1970s, http://www.psp.wa.gov/downloads/ANS/MISM_Online.pdf; and in the sampling sites not earlier than the late 1980s, B. Rinkevich, pers. observ.). Thus, the significant differences between the northern, younger

populations and the southern, more established populations could be explained by different source genotypes arriving through differing maritime trajectories, by the high mutation rates (Reem et al., 2013a) or by multiple catastrophic events. From two clusters ($K = 2$), one of them ascribes to the Washingtonian populations, but on the other hand, according to $K = 5$ the two Washingtonian populations also separated from each other. mtDNA result also shows different haplotype groups for two very close Washingtonian populations (SM in H-2 and DMM in H-1). These results further attest to the existence of different assemblages of *B. schlosseri* genotypes in the USA Pacific populations, less than a century after their initial establishment.

5. Conclusions

In this study we characterized the population genetic fluctuations of *B. schlosseri* from 5 microsatellite loci in the Moss Landing, CA area and from the COI region throughout a period of 12 years and seven sampling session, and our results were supplemented by sampling (during 2001) five additional West Coast *B. schlosseri* populations along a 1520 km coastal stretch (from the Seattle area, WA, in the north to Santa Barbara, CA, in the south). This enabled us to further evaluate the impacts of catastrophic fresh-water flood events on genetic differentiation, diversity and population structures in the persistent, but transient, *Botryllus schlosseri* population from Moss Landing, CA (the EYC population). The wide geographical range of the analyses has also allowed us to form a calculated estimation regarding the source of the EYC population, which was wiped out by the 2005 flood event. While the mtDNA results did not strongly support a pattern of genetic differentiation connected with the flood catastrophes, it provided valuable information regarding the six USA West Coast populations. On the other hand, the neutral microsatellite markers revealed strong genetic associations with catastrophic events. Significant overall θ, G_{ST} and D_{est} values, significant pairwise G_{ST} and D_{est} values in the 2001–2007 samples, allelic fluctuation and private alleles, all revealed connections between fresh-water flood catastrophes and population genetic characteristics. Based on the present and previous studies results (Reem et al., 2013a), we conclude that the SCH population is one of the most probable recreation source of the EYC population though the possibility of contribution from faraway populations cannot be ignored.

Competing interest

The authors have decelerated that no competing interests exist.

Acknowledgements

This study was supported by Grants from the Israeli-Academy-of-Science (68/10) and PERSEUS-FP7. Thanks to Kenan Murat Karahan for drawing Fig. 1 and to Alison Kideys and Lilach Baumer for copy-editing the manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2016.05.005.

References

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48. http://dx.doi.org/10.1093/oxfordjournals.molbev.a026036.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F., 1996. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Lab. génome, Popul. Interact. CNRS UMR 5000, 1996–2004.

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. http://dx.doi.org/10.1007/s10021-006-0141-y.

Ben-Shlomo, R., Motro, U., Paz, G., Rinkevich, B., 2008. Pattern of settlement and natural chimerism in the colonial urochordate *Botryllus schlosseri*. Genetica 132, 51–58. http://dx.doi.org/10.1007/s10709-007-9148-3.

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

Bernier, R.Y., Locke, A., Hanson, J.M., 2009. Lobsters and crabs as potential vectors for tunicate dispersal in the southern Gulf of St. Lawrence, Canada. Aquat. Invasions 4, 105–110. http://dx.doi.org/10.3391/ai.2009.4.1.11.

Bock, D.G., MacIsaac, H.J., Cristescu, M.E., 2012. Multilocus genetic analyses differentiate between widespread and spatially restricted cryptic species in a model ascidian. Proc. R. Soc. B Biol. Sci. 279, 2377–2385. http://dx.doi.org/10.1098/rspb.2011.2610.

Chapuis, M.P., Estoup, A., 2007. Microsatellite null alleles and estimation of population differentiation. Mol. Biol. Evol. 24, 621–631. http://dx.doi.org/10.1093/molbev/msl191.

Caputi, L., Crocetta, F., Toscano, F., Sordino, P., Cirino, P., 2014. Long-term demographic and reproductive trends in *Ciona intestinalis* sp. A. Mar. Ecol. http://dx.doi.org/10.1111/maec.12125.

Chadwick-Furman, N.E., Weissman, I.L., 1995. Life histories and senescence of *Botryllus schlosseri* (Chordata, Ascidiacea) in Monterey Bay. Biol. Bull. 189, 36–41. http://dx.doi.org/10.2307/1542199.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B 39, 1–38. http://dx.doi.org/10.2307/2984875.

Earl, D.A., VonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361. http://dx.doi.org/10.1007/s12686-011-9548-7.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol. Ecol. 14, 2611–2620. http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. 1, 47–50.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299. http://dx.doi.org/10.1371/journal.pone.0013102.

Forster, P., Harding, R., Torroni, A., Bandelt, H.J., 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. Am. J. Hum. Genet. 59, 935–945.

Fu, Y.X., Li, W.H., 1993. Statistical tests of neutrality of mutations. Genetics 133, 693–709.

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

Goudet, J., 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. J. Hered. 86, 485–486. http://dx.doi.org/10.1371/journal.pone.0026694.

Goudet, J., 1999. PCA-GEN for Windows. University of Lausanne, Lausanne, Switzerland.

Grosberg, R., 1987. Limited dispersal and proximity-dependent mating success in the colonial ascidian *Botryllus schlosseri*. Evolution 41, 372–384. http://dx.doi.org/10.2307/2409145.

Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41, 95–98, 691774.

Harpending, H.C., 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum. Biol. 66, 591–600.

Harrison, X.A., Bearhop, S., Inger, R., Colhoun, K., Gudmundsson, G.A., Hodgson, D., McElwaine, G., Tregenza, T., 2011. Heterozygosity-fitness correlations in a migratory bird: an analysis of inbreeding and single-locus effects. Mol. Ecol. 20, 4786–4795. http://dx.doi.org/10.1111/j.1365-294X.2011.05283.x.

Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801–1806. http://dx.doi.org/10.1093/bioinformatics/btm233.

Jost, L., 2008. GST and its relatives do not measure differentiation. Mol. Ecol. 17, 4015–4026. http://dx.doi.org/10.1111/j.1365-294X.2008.03887.x.

Kalinowski, S.T., 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. Conserv. Genet. 5, 539–543. http://dx.doi.org/10.1023/B:COGE.0000041021.91777.1a.

Kalinowski, S.T., 2005. HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol. Ecol. Notes 5, 187–189. http://dx.doi.org/10.1111/j.1471-8286.2004.00845.x.

Kozloff, E.N., 1975. Light's manual: intertidal invertebrates of the central California coast. J. Fish. Res. Board. Canada 32, 2585–2586.

Lambert, G., 2001. A global overview of ascidian introductions and their possible impact on the endemic fauna. In: Sawada, H., Yokosawa, H., Lambert, C.C. (Eds.), The Biology of Ascidians. Springer-Verlag, Tokyo, 470 pp.

Lavaniegos, B.E., Ohman, M.D., 2003. Long-term changes in pelagic tunicates of the California current. Deep Sea Res. Part II Top. Stud. Oceanogr. 50, 2473–2498. http://dx.doi.org/10.1016/S0967-0645(03)00132-2.

Lejeusne, C., Bock, D.G., Therriault, T.W., MacIsaac, H.J., Cristescu, M.E., 2011. Comparative phylogeography of two colonial ascidians reveals contrasting invasion histories in North America. Biol. Invasions 13, 635–650. http://dx.doi.org/10.1007/s10530-010-9854-0.

Manni, L., Burighel, P., 2006. Common and divergent pathways in alternative developmental processes of ascidians. BioEssays 28, 902–912. http://dx.doi.org/10.1002/bies.20462.

Mantel, N., 1967. The detection of disease and a generalized regression approach. Cancer Res. 27, 209–220.

Nei, M., 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U.S.A. 70, 3321–3323.

Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89, 583–590.

Novak, S.J., 2007. The role of evolution in the invasion process. Proc. Natl. Acad. Sci. U.S.A. 104, 3671–3672. http://dx.doi.org/10.1073/pnas.0700224104.

Pancer, Z., Gershon, H., Rinkevich, B., 1994. Direct typing of polymorphic microsatellites in the colonial tunicate *Botryllus schlosseri* (Ascidiacea). Biochem. Biophys. Res. Commun. 203, 646–651. http://dx.doi.org/10.1006/bbrc.1994.2231.

Paz, G., Douek, J., Mo, C., Goren, M., Rinkevich, B., 2003. Genetic structure of *Botryllus schlosseri* (Tunicata) populations from the Mediterranean coast of Israel. Mar. Ecol. Prog. Ser. 250, 153–162. http://dx.doi.org/10.3354/meps250153.

Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics 28, 2537–2539. http://dx.doi.org/10.1093/bioinformatics/bts460.

Pearson, K., 1901. On lines and planes of closest fit to systems of points in space. Philos. Mag. J. Sci. 2, 559–572. http://dx.doi.org/10.1080/14786440109462720.

Pérez-Portela, R., Turon, X., Bishop, J.D.D., 2012. Bottlenecks and loss of genetic diversity: spatio-temporal patterns of genetic structure in an ascidian recently introduced in Europe. Mar. Ecol. Prog. Ser. 451, 93–105. http://dx.doi.org/10.3354/meps09560.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959. http://dx.doi.org/10.1111/j.1471-8286.2007.01758.x.

Reem, E., Douek, J., Katzir, G., Rinkevich, B., 2013a. Long-term population genetic structure of an invasive urochordate: the ascidian *Botryllus schlosseri*. Biol. Invasions 15, 225–241. http://dx.doi.org/10.1007/s10530-012-0281-2.

Reem, E., Mohanty, I., Katzir, G., Rinkevich, B., 2013b. Population genetic structure and modes of dispersal for the colonial ascidian *Botryllus schlosseri* along the Scandinavian Atlantic coasts. Mar. Ecol. Prog. Ser. 485, 143–154.

Rice, W.R., 1989. Analyzing tables of statistical tests. Evolution 43, 223–225. http://dx.doi.org/10.2307/2409177.

Rinkevich, B., Weissman, I.L., 1987. The fate of *Botryllus* (Ascidiacea) larvae cosettled with parental colonies: beneficial or deleterious consequences. Biol. Bull. 173, 474–488.

Rinkevich, B., Lauzon, R.J., Brown, B.W., Weissman, I.L., 1992. Evidence for a programmed life span in a colonial protochordate. Proc. Natl. Acad. Sci. U.S.A. 89, 3546–3550. http://dx.doi.org/10.1073/pnas.89.8.3546.

Rinkevich, B., Shapira, M., 1998. An improved diet for inland broodstock and the establishment of an inbred line from *Botryllus schlosseri*, a colonial sea squirt (Ascidiacea). Aquat. Living Resour. 11, 163–171. http://dx.doi.org/10.1016/S0990-7440(98)80113-7.

Rinkevich, Y., Paz, G., Rinkevich, B., Reshef, R., 2007. Systemic bud induction and retinoic acid signaling underlie whole body regeneration in the urochordate *Botrylloides leachi*. PLoS Biol. 5, 900–913. http://dx.doi.org/10.1371/journal.pbio.0050071.

Rinkevich, Y., Voskoboynik, A., Rosner, A., Rabinowitz, C., Paz, G., Oren, M., Douek, J., Alfassi, G., Moiseeva, E., Ishizuka, K.J., Palmeri, K.J., Weissman, I.L., Rinkevich, B., 2013. Repeated, long-term cycling of putative stem cells between niches in a basal chordate. Dev. Cell 24, 76–88. http://dx.doi.org/10.1016/j.devcel.2012.11.010.

Roman, J., Darling, J.A., 2007. Paradox lost: genetic diversity and the success of aquatic invasions. Trends Ecol. Evol. http://dx.doi.org/10.1016/j.tree.2007.07.002.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574. http://dx.doi.org/10.1093/bioinformatics/btg180.

Rosner, A., Moiseeva, E., Rinkevich, Y., Lapidot, Z., Rinkevich, B., 2009. Vasa and the germ line lineage in a colonial urochordate. Dev. Biol. 331, 113–128. http://dx.doi.org/10.1016/j.ydbio.2009.04.025.

Rousset, F., 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Resour. 8, 103–106. http://dx.doi.org/10.1111/j.1471-8286.2007.01931.x.

Rozas, J., Rozas, R., 1999. DnaSP version 5: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15, 174–175. http://dx.doi.org/10.1093/bioinformatics/15.2.174.

Rozas, J., Sánchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497. http://dx.doi.org/10.1093/bioinformatics/btg359.

Ruiz, G.M., Fofonoff, P.W., Carlton, J.T., Wonham, M.J., Hines, A.H., 2000. Invasion of coastal marine communities in North America: apparent patterns, processes, and biases. Annu. Rev. Ecol. Syst. 31, 481–531. http://dx.doi.org/10.1146/annurev.ecolsys.31.1.481.

Scofield, V.L., Schlumpberger, J.M., West, L.A., Weissman, I.L., 1982. Protochordate allorecognition is controlled by a MHC-like gene system. Nature 295, 499–502. http://dx.doi.org/10.1038/295499a0.

Saillard, J., Forster, P., Lynnerup, N., Bandelt, H.J., Nørby, S., 2000. mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. Am. J. Hum. Genet. 67, 718–726. http://dx.doi.org/10.1086/303038.

Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139, 457–462.

Stoner, D.S., Quattro, J.M., Weissman, I.L., 1997. Highly polymorphic microsatellite loci in the colonial ascidian *Botryllus schlosseri*. Mol. Mar. Biol. Biotechnol. 6, 163–171.

Stoner, D.S., Ben-Shlomo, R., Rinkevich, B., Weissman, I.L., 2002. Genetic variability of *Botryllus schlosseri* invasions to the east and west coasts of the USA. Mar. Ecol. Prog. Ser. 243, 93–100.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Tiozzo, S., Ballarin, L., Burighel, P., Zaniolo, G., 2006. Programmed cell death in vegetative development: apoptosis during the colonial life cycle of the ascidian *Botryllus schlosseri*. Tissue Cell 38, 193–201. http://dx.doi.org/10.1016/j.tice.2006.02.003.

United States Navy, 1951. Report on Marine Borers and Fouling Organisms in 56 Important Harbors and Tabular Summaries of Marine Borer Data from 160 Widespread Locations. Bureau of Yards and Docks, Dept. Navy, Washington, DC, 326 p.

Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58, 236–244. http://dx.doi.org/10.1080/01621459.1963.10500845.

Weir, B.S., 1979. Inferences about linkage disequilibrium. Biometrics 35, 235–254. http://dx.doi.org/10.2307/2529947.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370. http://dx.doi.org/10.2307/2408641.

Wright, S., 1931. Evolution in Mendelian populations. Genetics 16, 97–159.

Yund, P.O., Feldgarden, M., 1992. Rapid proliferation of historecognition alleles in populations of a colonial Ascidian. J. Exp. Zool. 263, 442–452. http://dx.doi.org/10.1002/jez.1402630412.