Genetic divergence and assortative mating between colour morphs of the sea urchin *Paracentrotus gaimardi*

I. Calderón[a], C.R.R.Ventura[b], X. Turon[c] and H.A. Lessios[d]

[a] Department of Animal Biology, Faculty of Biology, University of Barcelona, 645 Diagonal Ave, 08028 Barcelona, Spain.

[b] Museu Nacional/UFRJ - Dept. Invertebrados Lab. Echinodermata. Quinta da Boa Vista s/n. São Cristóvão. Rio de Janeiro, Brasil.

[c] Center for Advanced Studies of Blanes (CEAB, CSIC), Accés a la Cala S. Francesc 14, 17300 Blanes (Girona), Spain.

[d] Smithsonian Tropical Research Institute, Balboa, Republic of Panama.

Keywords: *Paracentrotus*, colour polymorphism, assortative mating, mitochondrial ATPase, histone, bindin.

Running title: Divergence between colour morphs in *P. gaimardi*

Corresponding author:
H.A. Lessios
Smithsonian Tropical Research Institute
Box 0843-03092
Balboa, Ancon
Republic of Panama
Email: LESSIOSH@si.edu
Telephone: +5072128253
Fax: + 5072128790

Abstract

Some species of sea urchins feature large variation in pigmentation. This variability may be the result of phenotypic plasticity, or it may be associated with genetic divergence between morphs. *Paracentrotus gaimardi* exhibits five colour morphs (pink, brown, green, grey and black), which often occur within the same locality. We studied genetic divergence between these morphs in three populations on the coast of Brazil. The entire mitochondrial ATPase 8-6 mitochondrial genes, a fragment of the intron of a nuclear histone and the entire nuclear gene coding for the sperm protein bindin were analyzed. Mitochondrial DNA was differentiated between the pink and all other morphs, but the histone intron was similar in all colour morphs. In bindin, nine codons were found to be under positive selection and significant differences of allelic frequencies were observed in almost all pairwise comparisons between colour morphs. Although the molecular differentiation in bindin is not large enough to suggest reproductive isolation, some degree of assortative mating within morphs seems to be occurring in this species.

Introduction

Diversity in colouration is a frequent phenomenon in marine invertebrates, but its ecological significance is often unknown. The relationship between genetic differentiation and pigmentation is complex, preventing general conclusions about the validity of using colour patterns to differentiate species (e.g., López-Legentil *et al*. 2003; Pérez-Portela *et al*. 2007; Pleijel *et al*. 2009). Many studies have attempted to understand the potential adaptive value of this colour plasticity, especially because discrete colour morphs may also vary in aspects other than appearance, such as behaviour, physiology or reproductive success. Colour variation has been used in taxonomic classification, but this trait may not reliably indicate evolutionary divergence. Colouration can evolve rapidly (Endler *et al*. 2005), outpacing other morphological characters.

In many species, differences in colour have been attributed to adaptation to local environments. Indeed, variations in colour may be related to age (Cuvier & Valenciennes 1828; Medioni *et al*. 2001), to light or wave exposure (Stoletzki & Schierwater 2005), to food availability or diet (Tlusty & Hyland 2005) or to behavioural patterns (Pryke 2007). Some species lack genetic differentiation between colour morphs (Verdyck *et al*. 1998) or feature variations in pigmentation not associated to existing genetic divergence (Sponer *et al*. 2001; Le Gac *et al*. 2004; Pérez-Portela *et al*. 2007; Pleijel *et al*. 2009). Alternatively, colour may correlate to genetic differences, often suggesting the existence of different lineages or species (e.g., Boury-Esnault *et al*. 1992; Allcock *et al*. 1998; Manchenko *et al*. 2000; Meroz-Fine *et al*. 2003; Tarjuelo *et al*. 2004; López-Legentil & Turon 2005; Hizi-Degany *et al*. 2007; Prada *et al*. 2008).

Colour variation in sea urchins is widespread (Millot 1964; Gras & Weber 1977; Growns & Ritz 1994; Coppard & Campbell 2004). Some species are capable of changing the intensity of their dermal colouration (Kleinholz 1938; Millot 1968; Jensen 1974), while others maintain the same colour throughout their life. The ecological implications of colour variation in echinoids are poorly known. In some species variation may be related to differences in habitat or behaviour (Lindahl & Runnström 1929; Tsuchiya & Nishihara 1984, 1985). The non-random distribution of colour morphs in Tasmanian populations of *Heliocidaris erythrogramma* suggests that this variation may be environmentally and/or genetically

controlled (Growns & Ritz 1994). In *Lytechinus variegatus*, breeding trials showed that colour is genetically controlled (Pawson & Miller 1982). Despite similarities in the gamete recognition protein bindin and in isozymes, subspecies *L. variegatus variegatus* and *L. variegatus carolinus* exhibit differences in colouration and in mitochondrial DNA (mtDNA), suggesting allopatric differentiation between populations (Zigler & Lessios 2004). Conversely, *L. variegatus atlanticus* from Bermuda, although differing in colour from the other two subspecies, shows no genetic differences from them. In the same genus, Cameron (1984) found that *L. anamesus* and *L. pictus*, off the coast of California and in the Sea of Cortez could be easily distinguished based on test colour; yet indiscriminate gamete association between the two morphs suggested that these two nominal species were merely ecotypes. Recent studies of bindin and mtDNA confirmed this conclusion (Zigler & Lessios 2004). Similarly, species of the genus *Echinometra* show a wide range of test and spinal colour variation. This variation is correlated with differences in mtDNA, bindin, morphology of tube foot spicules and size of the sperm (e.g., Metz *et al*. 1994; Metz & Palumbi 1996; Arakaki *et al*. 1998; Rahman *et al*. 2000; Landry *et al*. 2003; Rahman & Uehara 2004a,b; Rahman *et al*. 2004; Kinjo *et al*. 2006), but there is also colour polymorphism within many of the species. Thus, in *Echinometra* there is good evidence that colouration is correlated with separate specific status, even though colour alone is not sufficient for species discrimination.

The genus *Paracentrotus* is composed of two allopatric species. *P. lividus* inhabits the entire Mediterranean and the Eastern Atlantic from the British Islands to Morocco, including the islands of the Macaronesia. The range of *P. gaimardi* spans the coast of Brazil, from Rio de Janeiro to Santa Catarina and also the Atlantic coast of Africa, from the Gulf of Guinea to Angola (Mortensen 1943). Each of these species exhibits several colours morphs. The pigments responsible for colouration in *P. lividus* have been identified (Goodwin & Srisukh 1950; Lederer 1952) and a genetic basis of their variation has been found by Louise & Benard (1993); no such studies are available for *P. gaimardi*. Five colours (pink, brown, grey, green and black) can be distinguished in the latter species, all living in sympatry along the coast of Brazil. No obvious differences exist in other aspects of morphology, habitat preferences, light or wave exposure or diet between the different colour morphs of *P. gaimardi* (C.R.R. Ventura, pers. obs.). As sea urchins have no special organs for light reception and as fertilization is external, colour cannot be involved directly in mate choice. Therefore, if

assortative mating is occurring between colour morphs, it most likely arises from genetic divergence in other traits, possibly correlated with colour differentiation.

We studied genetic differentiation among the five sympatric colour morphs of *Paracentrotus gaimardi* based on three molecular markers: the mitochondrial ATPase 8-6 genes, a fragment of the intron of the nuclear histone, and the gene coding for the entire bindin molecule. Bindin is a sperm protein that plays an essential role in successful fertilization and may thus be directly responsible for non-random mating, eventually leading to speciation. The main aim of this study was to determine whether phenotypic colouration is correlated to molecular differentiation between morphs and, if so, how this differentiation affects mate choice in *P. gaimardi*.

Material and methods

Paracentrotus gaimardi was sampled at the coast of Brazil in three sites separated by 30 to 180 km (Fig. 1). Gonads from samples collected in Praia Vermelha were preserved in 70% ethanol, whereas gonads of samples from Itaipú and Prainha were preserved in absolute ethanol. Individuals from all 5 colour morphs of this species were collected from each location (Table 1).

Genomic DNA was extracted using the DNeasy Tissue kit (QIAGEN) and three molecular markers were analyzed. First, a fragment of the mitochondrial ATPase 8-6 region was amplified for 135 individuals with the primers LYSa (5'-AAG.CTT.TAA.ACT.CTT.AAT.TTA.AAA.G-3') and ATP6b (5'-GCC.AGG.TAG.AAC.CCG.AGA.AT-3'). Second, a fragment of the nuclear intron of histone, located between exons H3 and H2a was amplified for 89 individuals using primers H3aRint (5'-GTC.ACC.ATC.ATG.CCC.AAG.GAT.AT-3') and H2aRev (5'-GGT.GAA.CAA.GTC.CCA.CTG.GAA.ACT-3'). This intron is about 850 bp-long and includes two microsatellite motifs. We analyzed two fragments of this intron located between the exons and the beginning of the repeated motifs on both 5' and 3' ends of the intron. At the 5´end, we sequenced 10 bp of the exon and 210 bp of the intron up to the repeated motifs and on the 3' end we sequenced 179 bp of the intron, starting at the repeated motifs, and 50 bp of the coding exon. These two fragments were concatenated into a single fragment of 449 bp.

For both markers, ATPase 8-6 and histone, amplifications were performed in a total volume of 25 μL using 1.5 mM of $MgCl_2$, 0.4 μM of each primer and 1 U of MasterAmp *Tfl* DNA Polymerase (Epicentre Biotechnologies, Madison, WI). PCR amplicons were cleaned with Exonuclease I and Shrimp Alkaline Phosphatase (USB Corporation, Cleveland, OH) and labelled with BigDye® Terminator v.3.1 (Applied Biosystems), with the same primers used for amplification, before being run in an ABI 3130 automated sequencer (Applied Biosystems). The gene coding for the entire mature bindin was also analyzed for 26 individuals from Praia Vermelha including representatives of all colour morphs (Table 1). This gene was amplified using primers PreCleF: (5'-CAG.GTG.ATA.CAG.AAA.GAA.GCG.GT-3') and Mid3'UTR (5'-ATG.TCG.TTG.CAA.TCA.TGA.AGG-3') designed for *Paracentrotus lividus* following the protocol described in Calderón *et al*. (2009). PCR products were sequenced directly using these same primers and, when heterozygotes were detected (either by differences in sequence length due to indels or to the presence of double peaks), products were cloned with pGEM-Easy Vector cloning kit (Promega) to separate individual alleles. Additionally, 8 presumed homozygote individuals were cloned in order to determine the reliability of direct sequencing in the assessment of homozygosity. Bacterial colonies were directly used as a template for PCR amplification with universal vector primers M13 F and M13 R following the protocol described in Calderón *et al*. (2009). Nucleotide differences that appeared only once in the whole data set were attributed to cloning artefacts and were replaced so as to match the consensus sequence (Villablanca *et al*. 1998; Calderón *et al*. 2009). For data analysis, homozygote individuals were considered as having two identical alleles.

Sequence analyses

In order to deduce the allelic phase of the histone sequences, we used the algorithms provided in PHASE v2.1 (Stephens *et al*. 2001; Stephens & Scheet 2005) as implemented in DnaSP 4.50 (Rozas *et al*. 2003). Haplotype and nucleotide diversity were calculated with DnaSP for all three markers. We used Arlequin ver. 3.1 (Excoffier *et al*. 2005) to calculate F_{ST} statistics based both on pairwise sequence differentiation and on allele frequencies between colour morphs and localities. The significance of the pairwise values was assessed by performing 10000 permutations, and *P*-values were corrected for multiple comparisons based on the false

discovery rate (FDR) control following Benjamini & Yekutieli's (2001) method (B-Y), according to which the critical value is determined by:

$$\alpha / \sum_{i=1}^{k} (1/i)$$

,

where k is the number of hypothesis tests performed and α is the experiment-wise error rate sought (Narum 2006) which we set at 0.05.

We used Arlequin for Analysis of Molecular Variance (AMOVA) to test hierarchical models of genetic variability based on allelic frequencies for ATPase and nuclear histone. Data were tabulated according to two categories: colour morph and geographical location. These data were used to estimate the relative contribution of these two grouping variables to the overall genetic variability. Additionally, given the direct implication of bindin in fertilization success in sea urchins, F_{ST} differentiation was computed based on both nucleotide sequences and on predicted amino acid sequences. For these analyses, bindin was translated using the universal nuclear code with BioEdit v.7.0.9.0 (Hall 1999). Finally, given the high excess of homozygotes observed in bindin (see below), we also used Arlequin to perform exact tests of Hardy-Weinberg equilibrium with 100000 steps in a Markov chain.

Tests for positive selection on bindin

We used MEGA 4.0 (Tamura *et al*. 2007) to calculate the proportion of non-synonymous (d_N) and synonymous (d_S) sites by the Pamilo & Bianchi (1993) and Li (1993) method for the whole bindin molecule. Values of ω ($=d_N/d_S$) greater than 1 are considered as evidence of positive selection (Zhang *et al*. 1997). We used the Recombination Detection Program (RDP2; Martin *et al*. 2005) to test for recombination in our sequences.

A Maximum-Likelihood (ML) tree was constructed from the bindin sequences of *Paracentrotus gaimardi* using TreeFinder (10000 replicates; Jobb 2007), based on the best-fit model of molecular evolution as determined by ModelTest v3.8 (Posada & Crandall 1998; Posada 2006) using the AIC criterion (Akaike 1974). This tree was used to determine which positions were subject to positive selection in this species, using the site models implemented in CodeML (PAML; Yang 2007). Models allowing and not allowing positive selection were compared using a likelihood ratio test [LRT; $2(lnL_2\text{-}lnL_1)$]. This was compared to the χ^2

distribution with degrees of freedom equal to the difference in the number of parameters between the general and nested model (Nielsen & Yang 1998; Yang *et al*. 2000). The models tested were the nearly neutral model M1a, which was compared to the positive selection model M2a, and the beta model M7, which was compared to the beta&ω model M8 (Swanson *et al*. 2001). These comparisons provide tests for positive selection (Yang *et al*. 2000). Briefly, both M1a and M2a specify two d_N/d_S ratios ($0<\omega_0<1$ and $\omega_1=1$), and M2a includes an additional site class with a d_N/d_S ratio $\omega_2 > 1$ estimated from the data. Similarly, M7 and M8 present a flexible beta distribution of d_N/d_S among sites but limited to the interval (0, 1), where 0 represents complete constraint and 1 represents the expectation of no selective constraint. Additionally, M8 considers an extra class of sites with $\omega_2>1$ estimated from the data (see Yang *et al*. 2005 for a detailed description of the terminology). Bayes Empirical Bayes (BEB) calculation of posterior probabilities for site classes was implemented for models M2a and M8 (Yang *et al*. 2005) to define sites subject to positive selection.

Finally, we constructed another ML tree using the bindin sequences of *Paracentrotus gaimardi* obtained in this study and the sequences of *P. lividus* previously analyzed by Calderón *et al*. (2009) to detect positive selection along the branches of each species. We used a branch-specific model with one ratio d_N/d_S for *P. gaimardi* (foreground branch) and another for *P. lividus* (background branch). We also used branch-sites models, which attempt to detect positive selection that affects only a few sites along each lineage. We computed the new branch-site test or test 2 of positive selection (Yang *et al*. 2005; Zhang *et al*. 2005), where model A with ω_2 estimated from the data is compared to the null model A with fixed $\omega_2=1$ (Yang *et al*. 2005; Zhang *et al*. 2005). This is a robust test for detecting positive selection on foreground branches.

Results

Diversity and population structure based on ATPase 8-6

We analyzed 532 bp of the gene coding for ATPase 8 and ATPase 6 from 135 individuals from all five colour morphs of *Paracentrotus gaimardi* (Table 1). No changes in sequence length were detected. Thirty-eight unique haplotypes were observed, of which 28 (74%) were singletons. One haplotype was present in more than 47% of the individuals analyzed, with all other haplotypes being represented at very low frequencies. Haplotype diversity of the entire

sample was high (H=0.762±0.038) and nucleotide diversity was very low (π=0.00315±0.00036), with an average number of nucleotide differences of 1.675. When haplotype diversities were computed individually for each colour morph, pink showed a slightly higher H than each of the other morphs, whereas grey showed a slightly smaller H (Fig. 2). Forty polymorphic sites were observed, of which 24 were parsimony informative and 12 represented non-synonymous changes.

F_{ST} values based on nucleotide pairwise differentiation showed no differences between colour morphs, pooling samples across localities (Table 2). However, when F_{ST} was calculated from haplotype frequencies, those of the pink morph were significantly different from frequencies of all other morphs, the highest differentiation being between pink and grey morphs (Table 2).

The AMOVA analyses revealed low but significant variation among colours (3.64%) and lower and non-significant variation among localities (2.59%). In the two analyses, most of the variation detected was found between individuals within colours or within localities, respectively, although the fixation indices were not significant (Table 3).

Diversity and population structure based on nuclear histone

Out of 89 individuals analyzed, PHASE revealed 47 alleles. Of these, two alleles were found in ca. 40% of the total sample (24.7 and 15.7% respectively), with each of the other alleles occurring in less than 7% of the sample. Twenty-seven individuals (31%) were homozygotes and 60 (69%) were heterozygotes. The analyzed fragment contained 39 polymorphic sites, with one synonymous substitution in the exons. Haplotype diversity (H=0.893±0.014) of the whole fragment was higher than that of ATPase, and nucleotide diversity was low (π=0.0046±0.0003). The non-coding fragment alone also displayed high haplotype diversity (H=0.890±0.015) and low nucleotide diversity (π=0.0053±0.0003). In contrast to ATPase, all morphs showed very similar H values (Fig. 3). The average number of nucleotide differences for the whole fragment was 2.083 (2.072 for the non-coding fragment).

F_{ST} analysis of population differentiation showed two significant differences between morphs based on nucleotide differentiation and two significant differences based on allelic frequencies. Pairwise F_{ST} values were small, and the significant differences indicated by the two methods of analysis were not consistent among colour morphs (Table 4).

Hierarchical analyses of molecular variance detected small and non-significant percentages of variation related to groupings by colour morphs or by geographic locations (1.37 and 0.49% respectively; Table 5). There was a modest but significant (2.5%) component of the variance associated to differentiation among colours within locations (Table 5). In both analyses, variation within locations or within colours, respectively, explained most of the variance (>95%, both being significant).

Bindin diversity and pattern of evolution

We analyzed 26 individuals of *Paracentrotus gaimardi* belonging to the 5 colour morphs of this species, all from a single locality (Table 1). Sequence lengths varied between 738 and 777 bp, coding for a protein of 246 to 259 amino acids from the cleavage site from preprobindin (RMKR) to the stop codon. As reported by Calderón *et al*. (2009) about bindin of *P. lividus*, bindin in *P. gaimardi* contains an intron located at the same conserved position as in other echinoid genera studied to date (Zigler & Lessios 2003). Program RDP2 did not find evidence of any recombination event in the sequence of *P. gaimardi* bindin.

We detected a total of 31 alleles in 34 sequences. Haplotype diversity was high in the whole sample (average H=0.991±0.001), and in each colour morph individually, and nucleotide diversity was low (average π=0.0125±0.0009; Fig. 4). We detected a total of 29 amino acid sequences in the 26 individuals analyzed. Eighteen individuals (69%) were homozygotes and eight (31%) were heterozygotes. Only one of the 8 homozygotes presumed from direct sequencing and subsequently cloned (12.5%) was in fact an heterozygote, suggesting that, although the percentage of homozygotes may be slightly overestimated in our results, cloning all presumed homozygotes would have been unlikely to have greatly altered our results. Exact tests showed that the excess of homozygotes over what is expected from Hardy-Weinberg equilibrium was significant for each colour morph (Table 6).

In bindin, F_{ST} values were computed for both nucleotide and amino acid sequences, using sequence differentiation and also allelic frequency differences. The significant pairwise comparisons detected for nucleotides were mirrored in the amino acid sequence analysis, so only the F_{ST} values corresponding to the former are shown (Table 7). Differentiation based on pairwise sequences revealed that all F_{ST} comparisons between green or brown to all other morphs (7 out of 10) were large and significant. F_{ST} values based on allelic frequencies were generally lower than values based on pairwise nucleotide differentiation, but all comparisons between colour morphs were significant (after correction), based both on nucleotide and on amino acid sequences.

The ML tree for bindin was constructed using the General Time Reversible model of evolution, with a proportion of invariable sites and a gamma correction (GTR+I+G: I= 0.8144; α=0.6213). The analysis for *Paracentrotus gaimardi* (Fig. 5) showed well-supported phylogenetic structure, but no relationship with colours. A ML tree including *P. gaimardi* and *P. lividus* showed that these species are reciprocally monophyletic with respect to bindin (data not shown). Neither the branch-specific nor the branch-site models based on a ML tree that included sequences from *P. gaimardi* and *P. lividus*, with the former as the foreground branch, detected any evidence of selection acting on *P. gaimardi* (Table 8).

The average proportion of synonymous and non-synonymous changes was similar in comparisons within and between morphs (Table 9). The ratio d_N/d_S for the whole bindin molecule was less than one (ω=0.033), indicating that positive selection was not acting on the entire molecule. In order to detect whether positive selection was acting on specific sites we used the CodeML option implemented in PAML applied to the ML tree. Models in PAML allowing for positive selection fit our data significantly better than alternative neutral models that did not allow positive selection (Table 8). Both models M2a and M8 identified the same nine sites as being subject to positive selection, with very high values of ω_2 (>12).

Discussion

The study of three markers in five sympatric colour morphs of *Paracentrotus gaimardi* from three populations along the coasts of Brazil yielded contrasting results. One colour morph

(pink) appeared clearly differentiated from all other morphs, especially the grey morph, based on a mitochondrial marker (Table 2). In the histone nuclear intron (Table 4), few differences among colours (not associated to any particular morph) were observed. The lower differentiation in the nuclear marker may be due to lower evolutionary rate relative to mitochondrial DNA; however, these differences may also indicate that inter-morph crosses are frequent enough to prevent allelic frequencies from drifting apart. In bindin (Table 7), a marker directly implicated in gamete recognition and under positive selective pressure, F_{ST} comparisons showed that there was significant divergence in allele frequencies between colour morphs. The finding of some significant differences between morphs in mitochondrial DNA and nuclear loci points to a non-random association between colour and genetic differentiation. The significant excess of homozygotes observed for bindin within each colour morph also indicates a lack of random mating between them.

The pattern of differentiation detected between morphs based on allelic frequencies generally differed from comparisons based on nucleotide sequences for all three markers (Tables 2, 4 and 7). It has been suggested that haplotype frequencies may be a better estimate of population differentiation than nucleotide pairwise comparisons when many very closely related haplotypes exist and little phylogenetic structure is present in the data. Differences in haplotype frequencies are established faster -at ecological time scales- than sequence divergence -which is established within evolutionary time-frames (i.e., O'Correy-Crowe *et al.* 1997; Fratini & Vannini 2002). Even though the genealogy of bindin of *Paracentrotus gaimardi* contains well-supported clades (Fig. 5), these clades do not match the colour composition of our samples. Thus, bindin divergence of colour morphs does not appear to be ancient enough to be reflected in evolutionary trees, but it nevertheless resulted in significant differentiation between all colour morphs as indicated by F_{ST} values based on allele frequencies

Previous studies have shown that departures from Hardy-Weinberg equilibrium in the form of deficit of heterozygotes are frequently observed in animals with planktonic dispersal, possibly associated to demography or mating systems (Addison & Hart 2005). The significant homozygote excess obtained for all colour morphs provides evidence of preferential intra-morph crosses. No-choice experimental crosses between the 5 colour morphs found that

fertilization success is high in all cases, but nevertheless higher in intra-morph crosses than in inter-morph crosses. Fertilization success is asymmetrical between morphs (C.R.R. Ventura, unpublished results). Additionally, post-fertilization factors such as differential survival or fitness of the resulting embryos, larvae and post-metamorphic stages may also contribute to the levels of differentiation found. Unfortunately, no information is available about mating preferences under conditions of mixed sperm, which in *Echinometra* have shown assortative mating not obvious in the no-choice experiments (Geyer & Palumbi 2005).

On the inter-specific level, various species of sea urchins show mating preferences at the gametic recognition stage (see review in Lessios 2007). On the intra-specific level, experimental studies in *Echinometra mathaei* have found that eggs exposed to sperm mixtures can select sperm from different males on the basis of the sperm's and the egg's bindin genotype (Palumbi 1999). In *Strongylocentrotus franciscanus*, sperm with different bindin genotypes showed different reproductive success, depending on their frequency and sperm density (Levitan & Ferrell 2006). Similar processes may be operating in *P. gaimardi*. These processes, however, may be too weak or of too recent origin to have caused reciprocal monophyly in bindin between colour morphs. They do, however, appear to have resulted in significant sequence and allelic frequency differences between morphs.

In conclusion, a small but measurable degree of assortative mating according to colour seems to exist in *Paracentrotus gaimardi*. The differentiation between colours may be recent and is best reflected in bindin allele frequency and sequence data. Further research would be necessary to ascertain whether bindin itself or some other linked trait is responsible for the observed genetic structure. Additional information about fertilization success between morphs and post-fertilization survival is also necessary to provide a more comprehensive view of the patterns of differentiation detected in this study.

Acknowledgements

We thank Ligia and Axel Calderón for their invaluable help with lab work. We are grateful to Laura Geyer for her help in obtaining permissions for sample importation. We are also grateful to Elinia M Lopes, Fernanda S. Viana and Mariana Contins for their help in the field and laboratory work in Brazil. This study was funded by project CTM2007-66635 of the

Spanish Government, by General Research Funds from STRI and by the Brazilian Council of Research (CNPq).

References

Addison JA, Hart MW (2005) Spawning, copulation and inbreeding coefficients in marine invertebrates. *Biology Letters*, **1**, 450-453.

Allcock AL, Watts PC, Thorpe JP (1998) Divergence of nematocysts in two colour morphs of the intertidal beadlet anemone *Actinia equina*. *Journal of the Marine Biological Association UK*, **78**, 821-828.

Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.

Arakaki Y, Uehara T, Fagoone I (1998) Comparative studies of the genus *Echinometra* from Okinawa and Mauritius. *Zoological Science*, **15**, 159-168.

Benjamini Y, Yekutieli D (2001) The control of false discovery rate under dependency. *Annual Statistics*, **29**, 1165-1188.

Boury-Esnault N, Solé-Cava AM, Thorpe JP (1992) Genetic and cytological divergence between colour morphs of the Mediterranean sponge *Oscarella lobularis* Schmidt (Porifera, Demospongiae, Oscarellidae). *Journal of Natural History*, **26**, 271-284.

Calderón I, Turon X, Lessios HA (2009) Characterization of the sperm molecule bindin in the sea urchin genus *Paracentrotus*. *Journal of Molecular Evolution*, **68**, 366-376.

Cameron RA (1984) Two species of *Lytechinus* (Toxopneustidae: Echinoidea: Echinodermata) are completely cross-fertile. *Bulletin of the Southern California Academy of Science*, **83**, 154-157.

Coppard SE, Campbell AC (2004) Taxonomic significance of spine morphology in the echinoid genera *Diadema* and *Echinothrix*. *Invertebrate Biology*, **123**, 357-371.

Cuvier G, Valenciennes A (1828) *Histoire Naturelles des Poissons*, Vol. III. Paris, Levrault.

Endler JA, Westcott DA, Madden JR, Robson T (2005) Animal visual systems and the evolution of color patterns: sensory processing illuminates signal evolution. *Evolution*, **59**, 1795-1818.

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

Fratini S, Vannini M (2002) Genetic differentiation in the mud crab *Scylla serrata* (Decapoda: Portunidae) within the Indian Ocean. *Journal of Experimental Marine Biology and Ecology*, **272**, 103-116.

Geyer LB, Palumbi SR (2005) Conspecific sperm precedence in two species of tropical sea urchins. *Evolution*, **59**, 97-105.

Goodwin TW, Srisukh S (1950) A study of the pigments of the sea urchins, *Echinus esculentus* L. and *Paracentrotus lividus* Lamarck. *Biochemistry Journal*, **47**, 69-76.

Gras H, Weber W (1977) Light-induced alterations in cell shape and ultrastructure of chromatophores of the sea urchin *Centrostephanus longispinus*. *Cell Tissue Research*, **182**, 167-176.

Growns JE, Ritz DA (1994) Colour variation in southern Tasmania populations of *Heliocidaris erythrogramma* (Echinometridae: Echinoidea). *Australian Journal of Marine and Freshwater Research*, **45**, 233-242.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95-98.

Hizi-Degany N, Meroz-Fine E, Shefer S, Ilan M (2007) Tale of two colors: *Cladopsammia gracilis* (Dendrophyliidae) color morphs distinguished by their genetics and ecology. *Marine Biology*, **151**, 2195-2206.

Jensen M (1974) The Strongylocentrotidae (Echinoidea), a morphological and systematic study. *Sarsia*, **57**, 113-148.

Jobb G (2007) *TREEFINDER version of February 2007*. Munich, Germany. Distributed by the author at www.treefinder.de.

Kinjo S, Uehara T, Yazaki T, Shirayama Y, Wada H (2006) Morphological diversity of larval skeletons in the sea urchin family Echinometridae (Echinoidea: Echinodermata). *Journal of the Marine Biological Association UK*, **86**, 799-816.

Kleinholz JH (1938) Color changes in echinoderms. *Pubblicazioni della Stazione Zoologica di Napoli* **17**, 53-57.

Landry C, Geyer LB, Arakaki Y, Uehara T, Palumbi SR (2003) Recent speciation in the Indo-West Pacific: rapid evolution of gamete recognition and sperm morphology in cryptic species of sea urchin. *Proceedings of the Royal Society of London, Series B*, **270**, 1839-1847.

Le Gac M, Féral JP, Poulin E, Veyret M, Chenuil A (2004) Identification of allopatric clades in the cosmopolitan ophiuroid species complex *Amphipholis squamata* (Echinodermata). The end of the paradox? *Marine Ecology Progress Series*, **278**, 171-178.

Lederer E (1952) Sur les pigments naphthoquinoniques des épines et du test des oursins *Paracentrotus lividus* and *Arbacia pustulosa*. *Biochimica et Biophysica Acta*, **9**, 92-101.

Lessios HA (2007) Reproductive isolation between species of sea urchins. *Bulletin of Marine Sciences*, **81**, 191-208.

Levitan DR, Ferrell DL (2006) Selection on gamete recognition proteins depends on sex, density and genotype frequency. *Science*, **312**, 267-269.

Li WH (1993) Unbiased estimation of the rates of synonymous and nonsynonymous substitutions. *Journal of Molecular Evolution*, **36**, 96-99.

Lindahl PE, Runnström R (1929) Variation und ökologie von *Psammechinus miliaris* (Gmelin). *Acta Zoologica* (Stockholm), **10**, 401-484.

López-Legentil S, Turon X (2005) How do morphotypes and chemotypes relate to genotypes? The colonial ascidian *Cystodytes* (Polycitoridae). *Zoologica Scripta*, **34**, 3-14.

López-Legentil S, Dieckmann R, Bontemps-Subielos N, Turon X, Banaigs B (2003) Chemical variation of alkaloids in color morphs of *Cystodytes* (Ascidiacea) *Biochemical Systematics and Ecology*, **33**, 1107-1119.

Louise F, Benard F (1993) Déterminisme génétique du polychromatisme de *Paracentrotus lividus* (Lamarck). *Bulletin de la Société Zoologique de France*, **118**, 405-408.

Manchenko GP, Dautova TN, Latypoy YY (2000) High level of genetic divergence between sympatric color morphs of the littoral sea anemone *Anthopleura orientalis* (Anthozoa: Actiniaria). *Biochemical Systematics and Ecology*, **28**, 737-750.

Martin DP, Williamson C, Posada D (2005) RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics*, **21**, 260-262.

Medioni E, Lecomte Finiger R, Louveiro N, Planes S (2001) Genetic and demographic variation among colour morphs of cabrilla sea bass. *Journal of Fish Biology*, **58**, 1113-1124.

Meroz-Fine E, Brickner I, Loya Y, Ilan M (2003) The hydrozoan coral *Millepora dichotoma*: speciation or phenotypic plasticity? *Marine Biology*, **143**, 1175-1183.

Metz EC, Kane RE, Yanagimachi H, Palumbi SR (1994) Fertilization between closely related sea urchins is blocked by incompatibilities during sperm-egg attachment and early stages of fusion. *Biological Bulletin*, **187**, 23-34.

Metz EC, Palumbi SR (1996) Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. *Molecular Biology and Evolution*, **13**, 397-406.

Millot N (1964) Pigmentary system of *Diadema antillarum* Philippi. *Nature*, **203**, 206-207.

Millot N (1968) *The enigmatic echinoids*. In: *Light as an ecological factor* (ed. Bainbridge R, Evans GC, Rackham O), pp 265-291. Blackwell Scientific. Oxford.

Mortensen T (1943) *A Monograph of the Echinoidea*. *Camarodonta*, Vol. III., C.A. Reitzel, Copenhagen.

Narum SR (2006) Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics*, **7**, 783-787.

Nielsen R, Yang Z (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*, **148**, 929-936.

O'Corry-Crowe GM, Suydam RS, Rosenberg A, Frost KJ, Dizon E (1997) Phylogeography, population structure and dispersal patterns of the beluga whale *Delphinapterus leucas* in the western Neartic revealed by mitochondrial DNA. *Molecular Ecology*, **6**, 955-970.

Palumbi SR (1999) All males are not created equal: fertility differences depend on gamete recognition polymorphisms in sea urchins. *Proceedings of the National Academy of Sciences USA*, **96**, 12632-12637.

Pamilo P, Bianchi O (1993) Evolution of the *Zfx* and *Zfy* genes: rates and interdependence between the genes. *Molecular Biology and Evolution*, **10**, 271-281.

Pawson DL, Miller JE (1982) *Studies of genetically controlled phenotypic characters in laboratory-reared* Lytechinus variegatus (*Lamarck) (Echinodermata: Echinoidea*). In: *Proceedings of the International Echinoderm Conference* (ed. Lawrence JM), pp. 165-171. Tampa Bay. AA Balkema, Rotterdam.

Pérez-Portela R, Duran S, Palacín C, Turon X (2007) The genus *Pynoclavella* (Ascidiacea) in the Atlanto-Mediterranean region: a combined molecular and morphological approach. *Invertebrate Systematics*, **21**, 187-205.

Pleijel F, Rouse G, Nygren A (2009) Five colour morphs and three new species of *Gyptis* (Hesionidae, Annelida) under a jetty in Edithburgh, South Australia. *Zoologica Scripta*, **38**, 89-99.

Posada D (2006) ModelTest Server: a web-based tool for the statistical selection of models of nucleotide substitution online. *Nucleic Acid Research*, **34**, 700-703.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817-818.

Prada C, Schizas NV, Yoshioka PM (2008) Phenotypic plasticity or speciation ? A case from a clonal marine organism. *BCM Evolutionary Biology*, **8**, 47-66.

Pryke SR (2007) Fiery red heads: female dominance among head color morphs in the Gouldian finch. *Behavioural Ecology*, **18**, 621-627.

Rahman MA, Uehara T (2004a) Interspecific and intraspecific variations in sibling species of sea urchin *Echinometra*. *Comparative Biochemistry and Physiology Part A*, **139**, 469-478.

Rahman MA, Uehara T (2004b) Interspecific hybridization and backcrosses between two sibling species of Pacific sea urchins (Genus *Echinometra*) on Okinawan Intertidal Reefs. *Zoological Studies*, **43**, 93-111.

Rahman MA, Uehara T, Aslan LM (2000) Comparative viability and growth of hybrids between two sympatric species of sea urchins (Genus *Echinometra*) in Okinawa. *Aquaculture*, **183**, 45-56.

Rahman MA, Uehara T, Pearse JS (2004) Experimental hybridization between two recently diverged species of tropical sea urchins, *Echinometra mathaei* and *Echinometra oblonga*. *Invertebrate Reproduction & Development*, **45**, 1-14.

Rozas J, Sánchez del Barrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496-2497.

Sponer R, Deheyn D, Roy MS (2001) Large genetic distances within a population of *Amphipholis squamata* (Echinodermata; Ophiuroidea) do not support colour varieties as sibling species. *Marine Ecology Progress Series*, **219**, 169-175.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *The American Journal of Human Genetics*, **68**, 978-989.

Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *The American Journal of Human Genetics*, **76**, 449-462.

Stoletzki N, Schierwater B (2005) Genetic and color morph differentiation in the Caribbean sea anemone *Condylactis gigantean*. *Marine Biology*, **147**, 747-754.

Swanson WJ, Yang Z, Wolfner MF, Aquadro CF (2001) Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proceedings of the National Academy of Science USA*, **98**, 2509-2514.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596-1599.

Tarjuelo I, Posada D, Crandall KA, Pascual M, Turon X (2004) Phylogeography and speciation of colour morphs in the colonial ascidian *Pseudodistoma crucigaster*. *Molecular Ecology*, **13**, 3125-3136.

Tlusty M, Hyland C (2005) Astaxanthin deposition in the cuticle of juvenile American lobster (*Homarus americanus*): implications for phenotypic and genotypic coloration. *Marine Biology*, **147**, 113-119.

Tsuchiya M, Nishihara M (1984) Ecological distribution of two types of sea urchin, *Echinometra mathaei* (Blainville), on Okinawan reef flat. *Galaxea*, **3**, 131-143.

Tsuchiya M, Nishihara M (1985) Agonistic behaviour and its effect on the dispersion pattern of two types of the sea urchin, *Echinometra mathaei* (Blainville). *Galaxea*, **4**, 37-48.

Verdyck P, de Wolf H, Desender K (1998) Genetic comparison of two colour morphs of *Phyllotreta tetrastigma* (Coleoptera: Chrisomelidae, Alticinae). *Belgian Journal of Zoology*, **128**, 119-134.

Villablanca FX, Roderick GK, Palumbi SR (1998) Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. *Molecular Ecology*, **7**, 547-560.

Yang Z (2007) PAML4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**, 1586-1591.

543 Yang Z, Nielsen R, Goldman N, Krabbe Pedersen AM (2000) Codon substitution models for heterogeneous
544 selection pressure at amino acid sites. *Genetics*, **155**, 431-449.
545 Yang Z, Wong WSW, Nielsen R (2005) Bayes Empirical Bayes inference of amino acid sites under positive
546 selection. *Molecular Biology and Evolution*, **22**, 1107-1118.
547 Zhang J, Kumar S, Nei M (1997) Small-sample tests of episodic evolution: a case study of primate lysozymes.
548 *Molecular Biology and Evolution*, **14**, 1335-1338.
549 Zhang J, Nielsen R, Yang Z (2005) Evaluation of an improved branch-site likelihood method for detecting
550 positive selection at the molecular level. *Molecular Biology and Evolution*, **22**, 2472-2479.
551 Zigler KS, Lessios HA (2003) 250 million years of bindin evolution. *The Biological Bulletin*, **205**, 8-15.
552 Zigler KS, Lessios HA (2004) Speciation on the coasts of the new world: phylogeography and the evolution of
553 bindin in the sea urchin genus *Lytechinus*. *Evolution*, **58**, 1225-1241.

Fig. 1. Sampled localities for *Paracentrotus gaimardi*. A) Praia Vermelha; B) Itapú; C) Prainha.

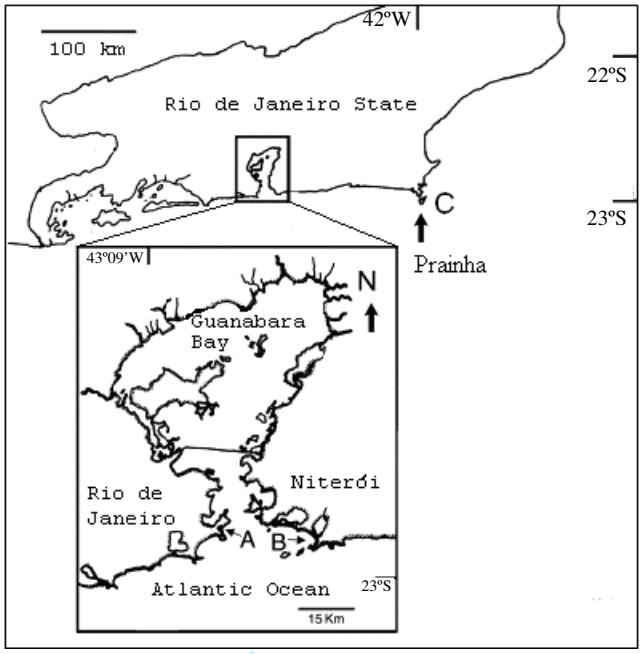

Fig. 2. Haplotype (grey bars) and nucleotide (black bars) diversity with one standard deviation for each morph based on mitochondrial ATPase 8 and 6.

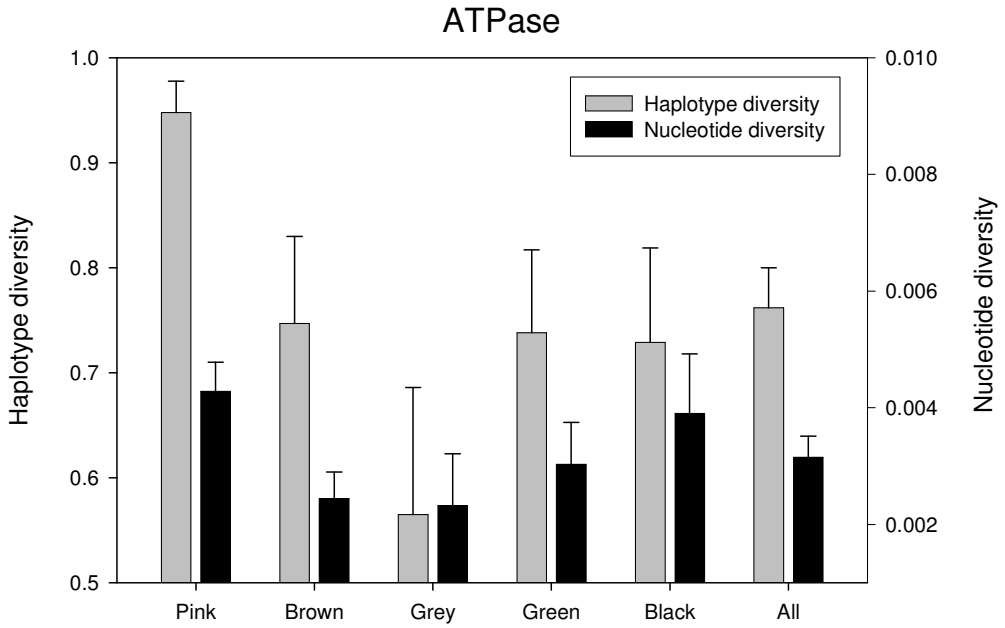

Fig. 3. Haplotype (grey bars) and nucleotide (black bars) diversity for histone with one standard deviation.

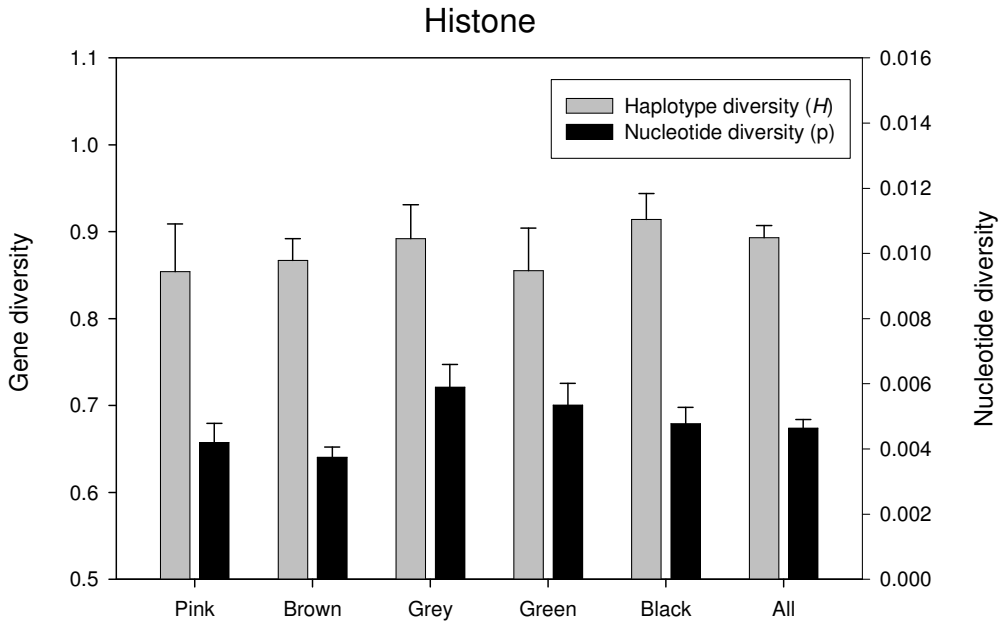

Fig. 4. Haplotype (grey bars) and nucleotide (black bars) diversity for bindin with one standard deviation.

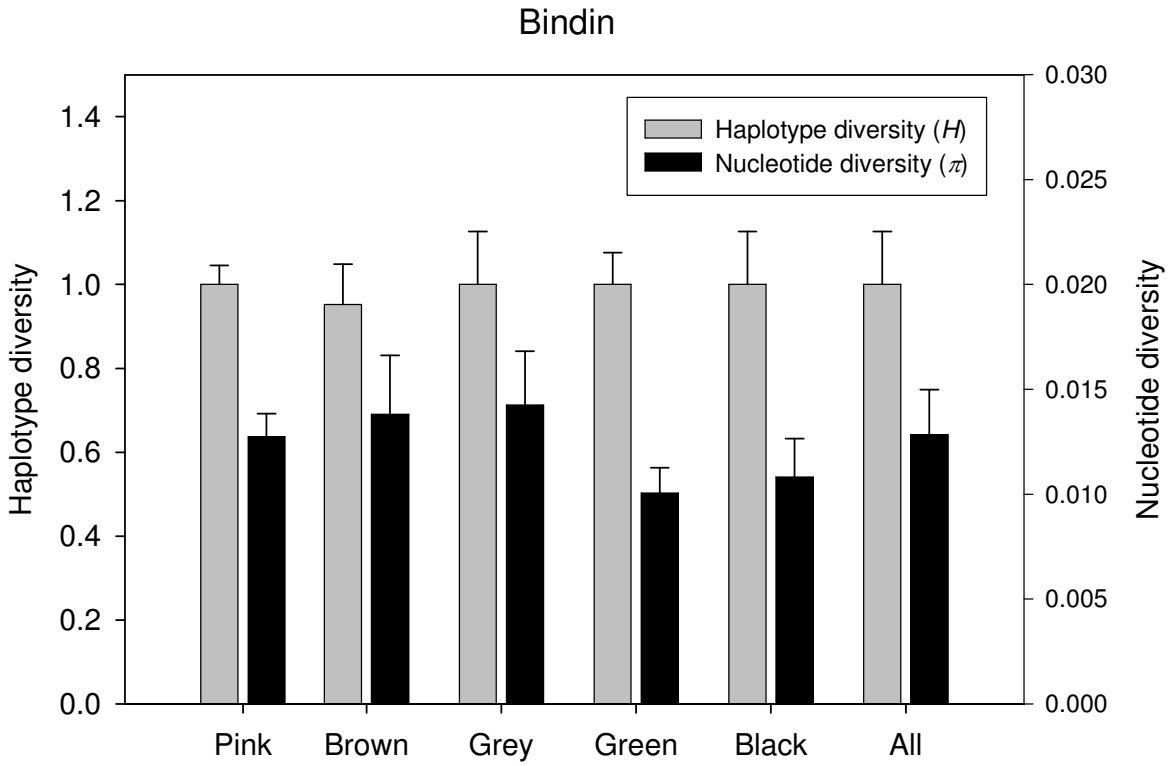

Fig. 5. ML tree of *Paracentrotus gaimardi* bindin. Individuals are coded according to their colour, numerals represent specimen number. Letters a and b represent two alleles in heterozygotes. Bootstrap support values are shown at the nodes when >50%.

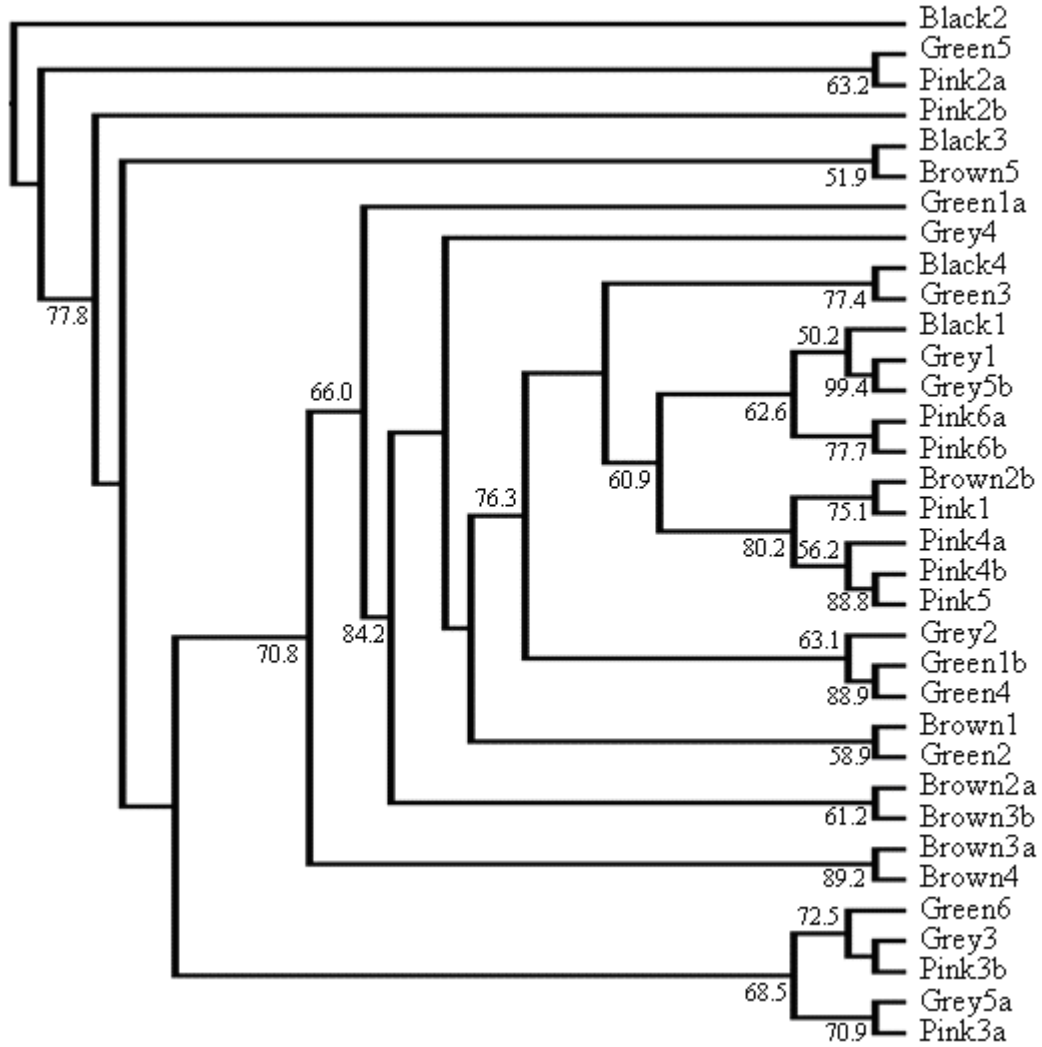

Table 1. Number of individuals per colour morph sampled at each location.

	Pink	Brown	Green	Grey	Black	Total
Itaipú	10/7	13/8	10/5	13/10	9/5	55/35
Prainha	4/2	9/9	12/7	6/6	13/9	44/33
Praia Vermelha	8/5/6	8/6/5	10/3/6	5/4/4	5/3/5	36/21/26
Total ATPase/Histone	22/14	30/23	32/15	24/20	27/17	135/89

First number in each cell indicates sample size analyzed for mitochondrial DNA, second number for nuclear histone, and third number (in Praia Vermelha) corresponds to individuals analyzed for bindin.

Table 2. F_{ST} for the analyzed fragment of mitochondrial ATPase between colour morphs.

Morphs	Pink	Brown	Grey	Green	Black
Pink		-0.0092	-0.0025	-0.0009	0.0025
Brown	0.0654*		0.0029	-0.0028	0.0233
Grey	0.1612*	0.0083		-0.0059	-0.0059
Green	0.0655*	-0.0139	-0.0076		-0.0024
Black	0.0794*	-0.0084	-0.0040	-0.0110	

Upper diagonal represents F_{ST} based on pairwise sequence differences whereas lower diagonal represents F_{ST} based on haplotype frequencies. *: comparisons significant after B-Y correction (Benjamini & Yekutieli 2001).

Table 3. AMOVA based on haplotype frequencies for ATPase.

AMOVA	df	% of variation	Fixation index	*P*
Among colours	4	3.64	F_{CT}: 0.0364	<0.05
Among locations within colours	10	-1.09	F_{SC}: -0.0113	0.7102
Within locations	120	97.45	F_{ST}: -0.0254	0.1219
Among locations	2	2.59	F_{CT}: -0.0075	0.5977
Among colours within locations	12	-0.75	F_{SC}: 0.0257	0.0948
Within colours	120	98.17	F_{ST}: 0.0183	0.1161

Two-level AMOVAs considering both colour morph and geographical location based on haplotype frequencies for ATPase.

Table 4. F_{ST} for the analyzed fragment of the intron of nuclear histone between colour morphs.

Morphs	Pink	Brown	Grey	Green	Black
Pink		0.0355*	-0.008	0.0749	0.0125
Brown	0.0294		0.0203*	0.0097	0.0023
Grey	-0.0131	0.0521		0.0584	0.0009
Green	0.0147*	-0.0052	0.0420*		0.0054
Black	-0.0056	0.0065	0.0203	-0.0043	

Upper diagonal represents F_{ST} based on pairwise sequence differences whereas lower diagonal represents F_{ST} based on allelic frequencies.*: comparisons significant after B-Y correction (Benjamini & Yekutieli 2001).

Table 5. AMOVA based on allelic frequencies for histone intron.

AMOVA	df	% of variation	Fixation index	*p*
Among colours	4	1.37	F_{CT}: 0.0137	0.1091
Among locations within colours	10	3.14	F_{SC}: 0.0168	0.1285
Within locations	163	95.80	F_{ST}: 0.0303	<0.05
Among locations	2	0.49	F_{CT}: 0.0049	0.242
Among colours within locations	12	2.50	F_{SC}: 0.0251	<0.05
Within colours	163	97.02	F_{ST}: 0.0298	<0.05

Two-level AMOVAs considering both geographical location and colour based on allelic frequencies for histone intron.

Table 6. Observed and expected heterozygosities for bindin with associated probabilities for deviations from Hardy-Weinberg equilibrium.

Morph	*Ho*	*He*	*P*
Pink	0.500	0.875	0.0035
Brown	0.400	0.800	0.0172
Green	0.167	0.847	0.0082
Grey	0.200	0.820	0.0004
Black	0.000	0.750	0.0009

Observed (*Ho*) and expected (*He*) heterozygosity for each colour morph. Probability (*P*) was calculated from an exact test.

Table 7. F_{ST} for bindin between colour morphs.

Morphs	Pink	Brown	Grey	Green	Black
Pink		0.087*	0.078	0.113*	0.024
Brown	*0.070**		0.143*	0.122*	0.124*
Grey	*0.067**	*0.109**		0.303*	0.023
Green	*0.053**	*0.093**	*0.091**		0.168*
Black	*0.070**	*0.111**	*0.109**	*0.093**	

Pairwise F_{ST} values in bindin between colour morphs calculated from nucleotide differentiation (above the diagonal) and allelic frequencies (below the diagonal). *: comparisons significant after B-Y correction (Benjamini & Yekutieli 2001).

Table 8. Comparison of models of variable $\omega=d_N/d_S$ ratios across sites with corresponding null models.

	L_1 (Neutral)	L_2 (Positive)	$2(lnL_2-lnL_1)$	df	*P*	Positively selected sites
Branch model (for *P. lividus* and *P. gaimardi*)						
Two-ratios *vs.* One-ratio	-2270.6235 (ω =0.4441)	-2270.6151 (ω_0=0.4310; ω_1= 0.4506)	0.0168	1	P=0.80	
Site model (for *P. gaimardi*)						
M1a vs M2a	-1802.5942 (p_0= 0.76142; p_1= 0.23858)	-1779.1287 (p_0=0.75836; p_1=0.21030; p_2=0.03134. ω_2=12.56446)	46.931	2	P<0.001	BEB (M2a): 11**; 33**; 42*; 48**; 51*; 63*; 64**; 214**; 273**.
M7 vs M8	-1802.9565 (p=0.0050; q=0.0197)	-1779.1373 (p_0=0.9682; p=0.0139; q=0.0468) (p_1=0.0318; ω_2=12.3763)	47.6384	2	P<0.001	BEB (M8): 11**; 33**; 42**; 48**; 51**; 63*; 64*; 214**; 273**.
Branch-site models (for *P. lividus* (as background) *vs. P. gaimardi*)						
Test 2 of positive selection (Model A with ω_1=1 *vs.* Model A with estimated ω_2)	-2242.22532 (p_0=0.5569; p_1=0.2333; p_2=0.1479; p_3=0.0619)	-2242.22533 (p_0=0.5589; p_1=0.2342; p_2=0.1457; p_3=0.0601) ω_2=1.0728	<0.001	1	P=0.95	

Likelihoods for each model and likelihood ratios for each comparison are shown. BEB: Bayes Empirical Bayes (Yang *et al*. 2005) is implemented in comparisons that involve models M2a and M8. *: P<0.05; **: P<0.01. Proportions (p or q) and ω are presented in parentheses for each model. Numbers identifying positively selected sites represent amino acid positions starting after the cleavage site.

Table 9. Replacement and silent substitutions in bindin within and between colour morphs.

			Pink	**0.008/0.022**	
		Brown	0.008/0.027	**0.008/0.024**	
	Grey	0.009/0.022	0.008/0.021	**0.007/0.016**	
Green	0.009/0.029	0.009/0.030	0.010/0.024	**0.010/0.024**	
Black	0.008/0.027	0.008/0.026	0.008/0.021	0.008/0.026	**0.007/0.020**
Morphs	Pink	Brown	Grey	Green	Black

Average number of amino acid replacement (d_N) substitutions per replacement site and silent (d_S) substitutions per silent site in bindin within and between colour morphs. First number corresponds to replacement and second number to silent changes. Values in bold along the diagonal correspond to intra-morph comparisons.