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

I. CALDERÓN,* C. R. R. VENTURA,† X. TURON‡ and H. A. LESSIOS§

*Department of Animal Biology, Faculty of Biology, University of Barcelona, 645 Diagonal Ave, 08028 Barcelona, Spain, †Museu Nacional/UFRJ – Dept. Invertebrados Lab. Echinodermata. Quinta da Boa Vista s/n, São Cristóào, Rio de Janeiro, Brazil, ‡Center for Advanced Studies of Blanes (CEAB, CSIC), Accés a la Cala S. Francesc 14, 17300 Blanes (Girona), Spain, §Smithsonian Tropical Research Institute, Balboa, Republic of Panama

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 side by side on the same rock. We studied genetic divergence between these morphs in three populations on the coast of Brazil. A fragment of the region encoding the mitochondrial ATPase 8 and 6 mitochondrial genes, a fragment of the intron of a nuclear histone and the entire nuclear gene coding for the sperm protein bindin were analysed. 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.

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

Received 18 September 2009; revision received 25 November 2009; accepted 29 November 2009

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. 2005; Pérez-Portela et al. 2007; Pleijel et al. 2009). Colour variation has been used in taxonomic classification, but this trait does not necessarily indicate evolutionary divergence, especially as colouration can evolve rapidly (Endler et al. 2005), outpacing other morphological characters. Variations in colour may be related to age (Cuvier & Valenciennes 1828; Medioni et al. 2001), to light or wave exposure (Stoletzki & Schierwater 2005), to diet (Tlusty & Hyland 2005) or to behavioural patterns (Pryke 2007). Some species lack genetic differentiation between colour morphs (e.g. 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; Alcock 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).

In sea urchins colour variation 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), whereas 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 nonrandom 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. More 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, 2004; Landry et al. 2003; Rahman & Uehara 2004a,b; 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. 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 colour 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 (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 sympathy 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, all of which can be found on the same rock (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 P. gaimardi based on three molecular markers: a fragment of the mitochondrial ATPase 8 and 6 genes, a fragment of the intron of the nuclear histone and the entire gene coding for the bindin molecule. Bindin is a sperm protein that plays an essential role in successful fertilization and may thus be directly responsible for nonrandom 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. Assortative mating by colour morphs can lead to incipient reproductive isolation with respect to colour, a phenomenon common among organisms that use visual cues in choosing their partners, such as birds and fishes, but not documented in invertebrates that are unable to form visual images.

**Materials and methods**

*Paracentrotus gaimardi* was sampled at the coast of Brazil in the state of Rio de Janeiro, at three sites separated by 30–150 km: Praiah a Arraial do Cabo (22°57' S, 42°00' W), Itaipu in Niterói (22°59' S, 43°04' W), and Praia Vermelha in Rio de Janeiro (22°57' S, 43°09' W). Individuals were collected from each location as encountered while snorkelling at water approximately 3 m deep. Initial collections were made without regard to colour, then additional individuals of a particular colour morph were added until 13 individuals from each of the five colour morphs of this species were collected from each location. Collections of the more common black morph were completed first and of the more rare pink morph were completed last. Gonads from samples collected in Praia Vermelha were preserved in 70% ethanol; gonads of samples from Itaipu and Praia de were preserved in absolute ethanol. Some individuals did not amplify well; the final sample sizes for each gene are shown in Table 1.

Genomic DNA was extracted using the DNeasy Tissue kit (QIAGEN), and three molecular markers were analysed. First, a fragment of the mitochondrial ATPase 8 and 6 region (532 bp) was amplified for 135 individuals with the primers LYSa (5'-AAG.CTT.TAA.ACT.CT- T.AAT.TTA.AAA.G-3') and ATP6b (5'-GCC.AGG.TA- G.AAC.CCG.AGA.AT-3'). Second, a fragment of the
Table 1 Number of individuals per colour morph sampled at each location

<table>
<thead>
<tr>
<th></th>
<th>Pink</th>
<th>Brown</th>
<th>Green</th>
<th>Grey</th>
<th>Black</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Itaiú</td>
<td>10/7</td>
<td>13/8</td>
<td>10/5</td>
<td>13/10</td>
<td>9/5</td>
<td>55/35</td>
</tr>
<tr>
<td>Prainha</td>
<td>4/2</td>
<td>9/9</td>
<td>12/7</td>
<td>6/6</td>
<td>13/9</td>
<td>44/33</td>
</tr>
<tr>
<td>Praia Vermelha</td>
<td>8/5/6</td>
<td>8/6/5</td>
<td>10/3/6</td>
<td>5/4/4</td>
<td>5/3/5</td>
<td>36/21/26</td>
</tr>
<tr>
<td>Total</td>
<td>22/14</td>
<td>30/23</td>
<td>32/15</td>
<td>24/20</td>
<td>27/17</td>
<td>135/89</td>
</tr>
</tbody>
</table>

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

nuclear intron of histone, located between exons H3 and H2a was amplified for 89 individuals using primers H3aRint (5′-GTC,ACC,ATG,CCC,AAG,GAT,AT-3′) and H2aRev (5′-GGA,CAA,GTC,CCA,CTG-GAA,ACT-3′). This intron is about 850 bp long and includes two microsatellite motifs. We analysed 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 178 bp of the intron, starting at the repeated motifs, and 51 bp of the coding exon. These two fragments were concatenated into a single fragment of 449 bp. For both markers, ATPase 8 and 6 and histone, amplifications were performed in a total volume of 25 μL using 1.5 mM of MgCl₂, 0.4 μM of each primer and 1 U of MasterAmp Tfl DNA Polymerase (Epicentre Biotechnologies). PCR amplicons were cleaned with Exonuclease I and Shrimp Alkaline Phosphatase (USB Corporation) and labelled with BigDye® Terminators 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 analysed 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 Mid3UTR (5′-ATG,TCG,TTG,CAA,TCA,TGA,AGG-3′), designed for Paracentrotus lividus, following the protocol described in Calderón et al. (2009). PCR products (up to 819 bp long) 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, eight 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. Sequences have been deposited in GenBank under accession numbers GU211850–GU211887 for ATPase8 and 6, GU211803–GU211849 for the Histone intron and GU211772–GU211802 for bindin.

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). We used Arlequin ver. 3.1 (Excoffier et al. 2005) to calculate FST 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 10 000 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:

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

where k is the number of hypothesis tests performed and \(\alpha\) 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, FST 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).

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
DIVERGENCE BETWEEN COLOUR MORPHS IN *P. GAIMARDI* 487

Pamilo & Bianchi (1993) and Li (1993) method for the whole bindin molecule. Values of \( \omega = \frac{d_{NS}}{d_S} \) 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 *P. gaimardi* using TreeFinder (10 000 replicates; Jobb 2007), based on the best-fit model of molecular evolution as determined by ModelTest v.3.8 (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; \( 2(\ln L_2 - \ln L_1) \)]. This was compared to the \( \chi^2 \) 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\&\& model M8 (Swanson et al. 2001). These comparisons provide tests for positive selection (Yang et al. 2000). Briefly, both M1a and M2a specify two \( d_{NS}/d_S \) ratios \( (0 < \omega_0 < 1 \text{ and } \omega_1 = 1) \), and M2a includes an additional site class with a \( d_{NS}/d_S \) ratio \( \omega_2 > 1 \) estimated from the data. Similarly, M7 and M8 present a flexible beta distribution of \( d_{NS}/d_S \) 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 determine sites subject to positive selection.

Finally, we constructed another ML tree using the bindin sequences of *P. gaimardi* obtained in this study and the sequences of *P. lividus* previously analysed by Calderón et al. (2009) to detect positive selection along the branches of each species. We used a branch-specific model with one \( d_{NS}/d_S \) ratio 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 a or test 2, of positive selection (Yang et al. 2005; Zhang et al. 2005), where model A with \( \omega_2 \) 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**

**Differences between colour morphs in ATPase 8 and 6**

We analysed 532 bp of the genes coding for ATPase 8 and ATPase 6 from 135 individuals from all five colour morphs of *P. gaimardi* (Table 1). Forty polymorphic sites were observed, of which 12 represented non-synonymous changes. 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 analysed, with all other haplotypes being represented at very low frequencies.

To find out whether the colour morphs were differentiated for ATPase8 and 6 we calculated \( F_{ST} \) statistics. \( F_{ST} \) 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). 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).

**Differences between colour morphs in nuclear histone.** Out of 89 individuals analysed, PHASE revealed 47 alleles. Of these, two alleles were found in ~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

<table>
<thead>
<tr>
<th>Morphs</th>
<th>Pink</th>
<th>Brown</th>
<th>Grey</th>
<th>Green</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( -0.0092 )</td>
<td>( -0.0025 )</td>
<td>( -0.0009 )</td>
<td>( 0.0025 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( 0.0654^* )</td>
<td>( 0.0029 )</td>
<td>( -0.0028 )</td>
<td>( 0.0233 )</td>
<td></td>
</tr>
<tr>
<td>Brown</td>
<td>( 0.1612^* )</td>
<td>( 0.0083 )</td>
<td>( -0.0059 )</td>
<td>( -0.0059 )</td>
<td></td>
</tr>
<tr>
<td>Grey</td>
<td>( 0.0655^* )</td>
<td>( -0.0139 )</td>
<td>( -0.0076 )</td>
<td>( -0.0024 )</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>( 0.0794^* )</td>
<td>( -0.0084 )</td>
<td>( -0.0040 )</td>
<td>( -0.0110 )</td>
<td></td>
</tr>
</tbody>
</table>

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).
and 60 (69%) were heterozygotes. The analysed fragment contained 39 polymorphic sites, with one synonymous substitution in the exons.

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). A modest (2.5%) but significant component of the variance was 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 significant).

**Bindin diversity and pattern of evolution.** We analysed 26 individuals of *Paracentrotus gaimardi* belonging to the five colour morphs of this species, all from a single locality (Table 1). Sequence lengths of mature bindin varied between 738 and 777 bp, coding for a protein of 246–259 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 sequences of *P. gaimardi* bindin.

There was a total of 31 unique alleles in 34 sequences and 29 amino acid sequences. Eighteen individuals (69%) were homozygotes and eight (31%) were heterozygotes. Only one of the eight 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.

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 6). Differentiation based on pairwise sequences revealed that all F_{ST} comparisons between green or brown to all other morphs (seven 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 of *P. gaimardi* 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; \( \alpha = 0.6213 \)). The analysis (Fig. 1) 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 (results 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 7).

The ratios of non-synonymous to synonymous changes were similar in comparisons within and between morphs (Table 8). The ratio \( \omega = d_{NS}/d_S \) for the whole bindin molecule was less than one (\( \omega = 0.033 \)), 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 Co-deML 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 7). Both models M2a and M8 identified the same nine sites as being subject to positive selection, with very high values of \( \omega_2 \) (>12).

Table 6

<table>
<thead>
<tr>
<th>Morphs</th>
<th>Pink</th>
<th>Brown</th>
<th>Grey</th>
<th>Green</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pink</td>
<td>0.087*</td>
<td>0.078</td>
<td>0.113*</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td>Brown</td>
<td>0.070*</td>
<td>0.143*</td>
<td>0.122*</td>
<td>0.124*</td>
<td></td>
</tr>
<tr>
<td>Grey</td>
<td>0.067*</td>
<td>0.109*</td>
<td>0.303*</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>0.053*</td>
<td>0.093*</td>
<td>0.091*</td>
<td>0.168*</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.070*</td>
<td>0.111*</td>
<td>0.109*</td>
<td>0.093*</td>
<td></td>
</tr>
</tbody>
</table>

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).

Discussion

The study of three markers in five sympatric colour morphs of \( P. gaimardi \) 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 6), 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 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 6). 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 in evolutionary time-frames (O’Corry-Crowe et al. 1997; Fratini & Vannini 2002). Even though the genealogy of bindin of *P. gaimardi* contains well-supported clades (Fig. 1), 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.

On the inter-specific level, various species of sea urchins show mating preferences at the gametic recognition stage (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

### Table 7 Comparison of models of variable $\omega = dN/dS$ ratios across sites with corresponding null models

<table>
<thead>
<tr>
<th>Branch model (for <em>P. lividus</em> and <em>P. gaimardi</em>)</th>
<th>L1 (Neutral)</th>
<th>L2 (Positive)</th>
<th>$2(\ln L2 - \ln L1)$</th>
<th>df</th>
<th>$P$</th>
<th>Positively selected sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-ratios vs. one-ratio</td>
<td>-2270.6235</td>
<td>-2270.6151</td>
<td>0.0168</td>
<td>1</td>
<td>$P = 0.80$</td>
<td></td>
</tr>
<tr>
<td>(\omega = 0.4441)</td>
<td>(\omega = 0.4310; \omega = 0.4506)</td>
<td>\n</td>
<td>Site model (for <em>P. gaimardi</em>)</td>
<td>-1802.5942</td>
<td>-1779.1287</td>
<td>46.931</td>
</tr>
<tr>
<td>M1a vs. M2a</td>
<td>(p = 0.76142; p = 0.23858)</td>
<td>(p = 0.75836; p = 0.21030; p = 0.03134. \omega = 12.56446)</td>
<td>\n</td>
<td>M7 vs. M8</td>
<td>-1802.9565</td>
<td>-1779.1373</td>
</tr>
</tbody>
</table>
| (P = 0.0050; q = 0.0197)                          | (p = 0.9682; P = 0.0139; q = 0.0468; \omega = 0.0318; \omega = 12.3763) | \n| Branch-site models (for *P. lividus* (as background) vs. *P. gaimardi*) | -2242.2532 | -2242.2533 | <0.001                | 1  | $P = 0.95$ | \n| Test 2 of positive selection                      | (p = 0.5569; p = 0.2333; p = 0.1479; p = 0.0619) | (p = 0.5589; p = 0.2342; p = 0.1457; p = 0.0601; \omega = 1.0728) | \n| (Model A with \omega = 1 vs. Model A with estimated \omega) | \n
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 \omega are presented in parentheses for each model. Numbers identifying positively selected sites represent amino acid positions starting after the cleavage site.

### Table 8 Ratio of replacement substitutions per replacement site to silent substitutions per silent site in bindin within (on the diagonal) and between (below the diagonal) colour morphs

<table>
<thead>
<tr>
<th>Morphs</th>
<th>Pink</th>
<th>Brown</th>
<th>Grey</th>
<th>Green</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pink</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown</td>
<td>0.30</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grey</td>
<td>0.41</td>
<td>0.38</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>0.31</td>
<td>0.30</td>
<td>0.42</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.30</td>
<td>0.31</td>
<td>0.38</td>
<td>0.31</td>
<td>0.35</td>
</tr>
</tbody>
</table>
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. Based on ATPase8 and δ6 divergence, calibrated by divergence in other sea urchin genera across the Isthmus of Panama, *P. gaimardi* and *P. lividus* speciated only half a million years ago (Calderon et al. 2009), so there may have not been sufficient time for bindin alleles to have sorted beyond the level of 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 *P. gaimardi*. The differentiation between colours may be recent and is best reflected in bindin allele frequency and sequence data. If the differentiation of bindin is, indeed, an on-going process heading towards higher divergence in reproductive molecules and more complete reproductive isolation, the colour morphs are on their way towards becoming new species through sympatric speciation. On the other hand, it is equally possible that weak assortative mating is counteracted by gene flow between colour morphs and recombination between bindin and genes that control colour. In this case, colour morphs will only turn into separate species if an extrinsic barrier separates them, an unlikely event, given their propensity to occupy the same habitat. Further research would be necessary to ascertain whether bindin itself or some other linked trait is responsible for the observed genetic structure. 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


Rahman MA, Uehara T, Pearse JS (2004) Experimental hybridization between two recently diverged species of


I. Calderón is interested in spatial and temporal small-scale genetic structure of populations. CRR Ventura is interested in phenotypic variation, reproduction, development and evolution of echinoderms. Xavier Turon works on biology of benthic invertebrates, including chemical ecology, reproductive biology and population genetics. H. Lessios is interested in speciation of marine organisms and the evolution of molecules involved in reproductive isolation.