Population Genetic Structure and Connectivity of the Abundant Sea Urchin, *Diadema setosum* around Unguja Island (Zanzibar)

Josefine Larsson, Oskar Henriksson and Mats Grahn

*School of Life Sciences, Södertörn University, BOX 4101 S-141 89, Huddinge Stockholm, Sweden.*

**Keywords:** sea urchin, *Diadema setosum*, Zanzibar, population genetics, AFLP.

**Abstract**—Uncontrolled growth of sea urchin populations may have a negative effect on coral reefs, making them barren. To avoid this, different methods of sea urchin reduction have been developed but, without knowledge of their genetic structure and connectivity, these methods may be ineffective. The aim of this study was to examine the fine-scale genetic structure and connectivity in the sea urchin, *Diadema setosum*, population around Unguja, Zanzibar, using AFLP. We found evidence of different genetic clusters, high migration between the sites and high genetic diversity within the sites. These findings indicate that a manual reduction of sea urchins with similar genetic connectivity, implemented on the same geographic scale as our study, would be ineffective since sites are probably repopulated from many sources.

**INTRODUCTION**

Coral reef ecosystems provide food and income for millions of people but, due to a number of mainly anthropogenic stressors, the resilience of coral reefs is diminishing (Nyström et al. 2000, Wilkinson 2004). One of the stressors is overfishing, as loss of predators may cause a shift in the ecological processes on a reef, changing both the intra- and inter-specific competitive interactions of prey organisms on the reef (Neudecker 1979, Wellington 1982, Hay et al. 1983, Hay 1984a, b, Lewis 1986, McClanahan & Muthiga 1989, McClanahan & Shafrir 1990). The effect of these changes may be illustrated by sea urchin populations that, through grazing, control the growth of macro-algae, maintaining the dominance of corals. However, due to extensive fishing of major sea urchin predators, e.g. trigger fish (Balistidae) and terminal-male wrasses (Labridae), populations of sea urchins are increasing (Hay 1984a and McClanahan & Muthiga 1988, McClanahan 1995a, 1998, 2000, Norström et al. 2009). Once sea urchins populations become too dense, they remove quantities of living and dead coral, degrading the coral reef and turning it into what is known as an urchin barren (McClanahan 1995a, 1998, McClanahan and Kurtis 1991, Carreiro-Silva & McClanahan 2001, McClanahan & Arthur 2001, Dumas et al. 2007, Norström et al. 2009). One of the most abundant sea urchins in East Africa that causes considerable bio-erosion is *Diadema setosum* (Leske) (Carreiro-
D. setosum is a broadcast spawner with external fertilization, high fecundity and a long larval duration; it can therefore spread over large distances (Strathmann 1978, Miller & Rarely 1999, Addison & Hart 2004, Cowen & Sponaugle 2009). Different strategies have been developed to prevent coral-dominated reefs being made barren by sea urchins; one involves lowering fishing pressure in areas with high sea urchin abundance and another is the “Sea Urchin Reduction” (SUR)-method where sea urchin populations are reduced by hand (McClanahan et al. 1996). The SUR-method can be improved if the genetic population structure is elucidated to detect migration and connectivity between populations (Pulliam 1988, Watkinson & Sutherland 1995), their genetic resilience (Nunes et al. 2009) and source/sink populations (Benzie & Wakeford 1997).

Different methods and markers are used in population genetic studies and allozyme and mitochondrial DNA (mtDNA) studies have been undertaken on D. setosum, revealing genetic differentiation on a large geographical and evolutionary scale (Lessios et al. 1996, Lessios & Pearse 1996, Lessios et al. 1998, Lessios et al. 1999, Lessios et al. 2001). The aim of this study was to investigate the fine-scale population genetic structure of the sea urchin D. setosum around Unguja, Zanzibar, Tanzania. In studies of this nature, high resolution markers are used but no microsatellites were available for D. setosum. However, Amplified Fragment Length Polymorphism (AFLP) has a resolution equal to or greater than microsatellites and no prior genetic information is needed, as non-species-specific primers are used to visualize genetic information in a presence/absence matrix (Bensch & Åkesson 2005). AFLP is a widely used method and is especially useful for differentiation between populations with either weak or strong genetic structure (Campell et al. 2003). A drawback with AFLP is that it only generates dominant marker data which makes evaluation of deviation from the Hardy-Weinberg equilibrium difficult, and it has limitations when it comes to the detection of immigration (Campell et al. 2003). However, as the aim of this study was to assess the genetic structure in a population and to establish whether recruitment comes from single or multiple sources, AFLP was the method of choice.

**MATERIAL AND METHODS**

**Study area**

Unguja Island, the largest island in the Zanzibar Archipelago, is located 35 km off the coast of Tanzania. The west coast of Unguja is protected from prevailing winds and strong currents, allowing the development of more diverse reefs than on the east coast (Johnstone et al. 1998). Four sites around Unguja Island were sampled: Chumbe Island (Ch), Bawe Island (Ba), Nungwi (Nu) and Jambiani (Ja) (Figure 1). All the sites were shallow, 0.5-3 m at low tide, their area ranging from 150 m² to 200 m², with urchin densities ranging from low (1-80) to medium (80-160) and high (>160). The urchin density was estimated by a rapid visual census with an estimate of their density. The sites ranged from a lush reef with high biodiversity (Chumbe) in a Marine Protected Area (MPA), to a degraded reef with almost no living coral (Nungwi) (Figure 1).

**Sampling**

A total of 203 sea urchins were collected with tongs at the four sites in November 2008 using SCUBA or snorkeling gear and mesh baskets. At each site, 49-55 individuals were randomly sampled within an area of similar habitat and divided into three size classes: small (20-39 mm), medium (40-49 mm) and large (>50 mm). The identification of each individual as *D. setosum* was confirmed according to the species-specific features of an orange ring around the anal entrance and five white spots evenly distributed around the body (Lessios & Pearse 1996).
Laboratory

Gonads were exposed by cutting across the tests (using scissor inserted through the peristomal membrane) and pulling the halves apart, exposing the gonad tissue. A piece of gonad was removed from each specimen, rinsed with double distilled water (DDW), dried on non-abrasive tissue paper and placed in extraction solution. Genomic DNA was isolated from a small piece of gonad tissue, following the protocol described by Laird et al. (1991). The DNA content in each sample was quantified in a NanoDrop© ND-1000 Spectrophotometer (NanoDrop technologies, Inc., USA) and diluted to a working concentration of 25 ng/µl. The samples were randomized before AFLP analysis to eliminate variations in the PCR reaction. AFLP-analysis was performed according to Vos et al. (1995), with modifications from Bensch et al. (2002). The urchin DNA (10µl of the 25 ng/µl extract) was incubated for 1 h at 37°C in a cocktail of 6.9 µl ddH₂O, 2 µl TA-buffer (10X), 1 µl BSA and 0.05 µl ECO R1 (Fermentas) (5’-G4AATTc-3’, 50 u/µl), 0.05 µl Tru/MseI (Fermentas) (5’-T4TAA-3’, 50 u/µl) in a total volume of 20 µl. After one hour of digestion, 5 µl ligation cocktail was added and the samples were incubated for 3 h at 37°C. The ligation cocktail contained 4.15 µl ddH₂O, 0.5 µl ligation buffer (10X), 0.025 µl E-adaptor (100 µM), 0.25 µl M-adaptor (corresponding to the fragments) and 0.1µl T4 Ligase (5 u/µl, Fermentas). The DNA was diluted with 180 µl ddH₂O and stored at -20°C. The pre-amplification was conducted with a Gene Amp ® PCR System 9700 PCR (Applied Biosystems) with the following temperature profile [94°C 2 min] + 94 °C -30 s, 56 °C -30 s, 72 °C -60 s] x 20 cycles + [72 °C - 10 min]. The DNA template (diluted, digested and ligated) was added to 10 µl of pre-amplification cocktail containing 1.8 µl ddH₂O, 2 µl MgCl (25mM), 2 µl PCR-buffer (10X), 4 µl dNTP (1mM), 0.06 µl
E-primer (5’-GACTGCGTACCAATTCC-3’), 0.06 µl M-primer (5’-GATGAGTCCTGAGTAGTA-3’) and 0.08 µl Taq polymerase (5 u/µl). The amplified samples were diluted with 180 µl ddH₂O and stored at -20ºC. Diluted, pre-amplified material (2.5 µl) was added to tubes containing 7.5 µl selective-amplification cocktail; 3.3 µl ddH₂O, 1 µl MgCl₂ (25mM), 1 µl PCR-buffer (10X), 2 µl dNTP (1mM), 0.06 µl FAM-labelled E-primer (5’-GACTGCGTACCAATTCCAT-3’), 0.06 µl M-primer(5’-GATGAGTCCTGAGTAGTA-3’) and 0.08 µl Taq polymerase (5 u/µl). The samples were incubated in a “touch-down” PCR: [94ºC-2min] + [94ºC-30 s, 65ºC - 0.7ºC/cycle-30 s, 72ºC-60 s] x 12 cycles + [94ºC-30 s, 56ºC-30 s, 72ºC-60 s] x 23 cycles +[72ºC-10 min]. One FAM-labelled primer combination was used (ECAT-MGTA), 18 duplicates being run at least twice. Selective-amplified material (3 µl) was added to a PCR-plate containing 87 µl ddH₂O and the DNA fragments were separated on an ABI3730XL capillary electrophoresis unit (Applied Biosystems) at Uppsala Genome Centre, Uppsala Sweden. The laboratory work and DNA extractions were conducted at the WIO-magnet facility at the Institute of Marine Science (IMS) in Stone town, Zanzibar. The AFLP analyses and data analyses were conducted at Södertörn University, Stockholm.

Data analysis

Data were scored in Genemapper 3.0 (Applied Biosystems). During this step, the unamplified samples were deleted and subsequent scoring was done using default AFLP settings with no normalization. The analysis range was set to 150-500 bp and the locus selection threshold was set to 200 rfu. The starting point of 150 bp was chosen to eliminate noise generated by primers. Genemapper 3.0 (Applied Biosystems) created a table based on PCR-fluorescence with peak heights for every locus. The table was exported into AFLP-score (Whitlock et al. 2008) to score genotype data from the AFLP markers (dominant molecular markers). This software interprets and normalizes the peak heights. In the program, the optimal scoring conditions are determined and used to create a genotype matrix used in further analysis. In addition, the program uses the normalized data together with the 18 duplicates to perform an error rate analysis (Whitlock et al. 2008), yielding, in this case, a phenotypic threshold of 20%, 200 rfu and 51% mismatch. The matrix obtained in AFLP-score was further analyzed in the computer software AFLP-surv to obtain values of genetic diversity (H) and Fst values of genetic diversity between the different sites (fixed geographic locations) (Vekemans et al. 2002). The data set was analysed assuming Hardy-Weinberg equilibrium using the Bayesian method with non-uniform prior distribution of allele frequencies (Zhivotovsky 1999). 1000 permutations were used to calculate the significance of the Fst. The AMOVA option implemented in Arlequin V. 3.11 (Excoffier et al. 2005) was used to test at which spatial scale the genetic differences occurred. STRUCTURE 2.2 was used to analyse the genetic structure of the populations and size classes. The burn was set to 50 000 with 50 000 additional cycles and the Bayesian approach was used. Each run was iterated three times. The program assigned each individual to a genetic cluster, depending on its genetic makeup, independent of the former information. Prior to the run, set up assumptions comprised an admixture of the populations with no correlation of allele frequencies. Clusters were created from the data, their number being calculated, with the K with the highest probability being indicated by the lowest Pr(X|K) (Pritchard et al. 2000). We explored a range from K=1 to K=20. ANOVA followed by a multiple comparison of size means of the STRUCTURE-generated clusters was performed to elucidate cohort recruitment.
RESULTS

Genetic variation

A total of 288 loci were detected by our primer combination, 105 which were polymorphic with a mean of 14.1 segregating loci per individual. The mean proportion of segregating loci per sampling site was between 40-50%. Sites were sampled in a nested fashion based on geographic distances to detect genetic structure at different spatial scales. The closest waterway distance was between Bawe and Chumbe Islands (15 km) and the longest distance was between Nungwi and Jambiani (155 km). The site-based analysis of AFLP revealed no significant genetic differences between the different sites (Fst -0.0008, p= 0.48). An analysis of molecular variance (AMOVA) was performed in Arlequin V. 3.11 (Excoffier et al. 2005) to test the hierarchical distribution of genetic variance between the nested sites. The sites were grouped into three groups based on their geographic distance, with Bawe and Chumbe in one group, and Jambiani and Nungwi in two separate groups. All the genetic variation was found within populations (Table 2) and no genetic structure could be found between the groups or sites. In the STRUCTURE analysis, the modal value for the distribution of ΔK was found where K= 6 (Fst 0.1234 p<0.05). The cluster distribution for each site is shown in Figure 2a. The mean assignment probabilities of individuals to clusters were high (0.9270, +/-0.127 SD, Fig. 2b). There was no correlation between geographic sites and genetic clusters, as all sites except Jambiani contained all clusters in more or less equal proportions (equal N in each cluster, Figure 2b.) Cluster number 4, the least numerous cluster, was not found at Jambiani. ANOVA followed by a multiple comparison of means was performed to determine the relationship between assigned clusters and the size of individual urchins (Figure 3); no significant relationships were found.

DISCUSSION

The results show a strong genetic structure but this was not based on geographic sampling sites as the Bayesian assignment test revealed that there were six genetic clusters at all the sites around Unguja Island but one (Jambiani lacked cluster 4). The presence of six significantly different clusters indicates a non-panmictic population structure and their wide distribution indicates a high level of connectivity between the sites. The result imply that the SUR method would thus be of limited effectiveness at most sites around Unguja island, as they would be re-populated from many sources. The high genetic variability within each site would also indicate a high level of genetic resilience towards stochastic events and disturbances (Miller et al. 2009, Nunes et al. 2009).

The ability of the sea urchin to rapidly re-colonize sites has been demonstrated in a SUR experiment by McClanahan et al. (1996), where the reduction plot had to be cleared every 1-3 months to keep it urchin-free. Their findings and our results both imply that SUR would be very labour intensive and expensive to carry out on populations with a high degree

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among groups</td>
<td>2</td>
<td>17.884</td>
<td>-0.06865 Va</td>
<td>-0.68</td>
</tr>
<tr>
<td>Within groups</td>
<td>1</td>
<td>13.189</td>
<td>0.06287 Vb</td>
<td>0.62</td>
</tr>
<tr>
<td>Within populations</td>
<td>199</td>
<td>2011.745</td>
<td>10.10927 Vc</td>
<td>100.06</td>
</tr>
<tr>
<td>Total</td>
<td>202</td>
<td>2042.818</td>
<td>10.1035</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Estimates of genetic differentiation (AMOVA) of AFLP- markers among and within groups and within populations of Diadema setosum around Unguja Island.
of genetic connectivity, such as found in this study. If a sea urchin population grows too large, they degrade a coral reef, turning it into an urchin barren. This can be considered an alternate steady state which is difficult to reverse as removing the stress does not automatically cause the ecosystem to return to its former state (Norström et al. 2009). Thus the SUR-method should only be conducted on protected reefs where there are enough herbivorous fish or sea urchins to graze on the algae; without this, uncontrolled algal growth might eventually dominate the reef (Carpenter 1990a, b, McClanahan 1995a, McClanahan et al. 1996, McClanahan 1998, Carreiro-Silva & McClanahan 2001, McClanahan & Arthur 2001, Dumas et al. 2007).

For a population’s genetic structure to be maintained there has to be a reproductive barrier between the different genetic clusters; separation by time or space (Grant et al. 1996), sperm recognition (Metz & Palumbi 1998) and microhabitat preferences (Bongaerts et al. 2010) are examples of such barriers. Strong genetic structuring was encountered in this study but the underlying reproductive barriers could not be discerned from our data. Neither site fidelity, nor cohort recruitment were the cause of the pattern as there was no correlation either between size and the genetic cluster, or sites and the genetic clusters. Other explanations may be that the genetic clusters have different microhabitat preferences, an attribute found in other marine organisms such as corals, where different genetic clusters have no physical barrier between them but occupy different habitats (Bongaerts et al. 2010). In this study, sampling was conducted at each site within similar habitat; however, microhabitat differences were not considered and may have been present. Further studies on microhabitat differences are therefore needed. Reproductive barriers can underlie cryptic speciation and, by using additional genetic markers in e.g. mtDNA, cryptic speciation may be revealed. In addition, biological studies on spawning time, spawning behaviour and sperm recognition might shed more light on the issue.

Fig. 2 a) Relative genetic cluster abundance of *Diadema setosum* within each site (n=number of individuals) based on AFLP markers derived from STRUCTURE; K= the number of clusters with the highest probability (Pritchard et al. 2000). b) Histogram indicating cluster assignment probabilities of *D. setosum*, samples derived from STRUCTURE.
Fig. 3 Plot of the mean size of *Diadema setosum* (test diameter, mm) within each cluster with 95% confidence intervals. An ANOVA followed by a pairwise comparison of means showed no significant relationship between the size of individual urchins and assigned cluster.

**Acknowledgments**—This work was supported by SIDA (as a Minor Field Study) and the Western Indian Ocean Marine Science Association (WIOMSA) through a Marine Science for Management (MASMA) grant. We thank Omri Bronstein, Dr. Sadri Said and the staff at the Institute of Marine Sciences, Zanzibar, for all help in facilitating the field excursion and laboratory work on Unguja. We are grateful to Mikael Lönn and anonymous reviewers for valuable comments on the manuscript.

**REFERENCES**


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

Grant PR, Grant RB, Deutsch JD (1996) Speciation and Hybridization in Island Birds. Philosophical Transactions: Biological Sciences 351: 765-772

Hay ME (1984a) Patterns of fish and urchin grazing on Caribbean coral reefs: are previous results typical? Ecology 65: 446-454


Richmond, MD (2002) A guide to the seashores of eastern Africa and the western Indian Ocean islands. SIDA/SAREC-USDM. Tanzania. 461 pp


