**Development of genetic tools for important African bioengineers: the case of the redbait species *Pyura herdmani* and *P. stolonifera* coastlines**

**A Dinoi1, M Rius1,2, M Tine1 and PR Teske1\***

*1 Centre for Ecological Genomics and Wildlife Conservation, Department of Zoology, University of Johannesburg, Johannesburg, South Africa*

*2 School of Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, United Kingdom*

*\* Corresponding author, email: pteske101@gmail.com*

Manuscript received July 2020 received September 2020; accepted December 2020

The development of new genomic resources is key for biological conservation in an era when the preservation of biodiversity is crucial. Here, we developed microsatellites and mitochondrial markers for the redbait species, *Pyura herdmani* and *P. stolonifera,* bothimportant bioengineers along African coastlines. Specifically, we identified microsatellites by means of pyrosequencing, and variable regions in the mitochondrial genome using RAD-seq libraries that were generated with Illumina sequencing. A total of 4 339 putative microsatellites were found, of which 32 were tested on 52 individuals collected along the south-east coast of South Africa. Of those, five microsatellites cross-amplified consistently in both species, and these were tested on a total of 260 samples. Additionally, we mapped the partial mitochondrial genome of both species and identified the most variable regions by comparing individuals from all regions where they have been recorded (including western South Africa and northwest Africa). The genetic markers developed here can be used to conduct population genetic studies, to monitor range shifts under contemporary climate change, and to help preserving ecosystem functioning along African coastlines.

**Keywords:** ascidians, Illumina sequencing, RAD-seq, microsatellites, mitogenome, STRs.

**Introduction**

It is widely accepted that the development of new molecular tools to investigate population genetics and phylogenetics can support biodiversity conservation actions (Goodwin et al. 2017). This is particularly true in cases where non-molecular methods alone are insufficient to readily distinguish between closely related species and/or morphologically conserved species, or to detect range shifts and hybridisation rates (Teske and Beheregaray 2009).

A widely used genetic marker for biodiversity conservation is the so-called microsatellites or short tandem repeats (STRs), which are codominantly inherited markers that are highly polymorphic (Weber and Wong 1993). They provide a tool to investigate small-scale resolution of demographic events and fine-scale population structure at a recent time-scale (Giles et al. 2018). Another invaluable genetic markers is the mitochondrial DNA (mtDNA), which is generally used to study the distribution of genetic variation across populations, and as a tool for reconstructing phylogenetic relationships between species (Griggio et al. 2014). However, its maternal inheritance creates a bias towards female-mediated processes (Avise 1994; Pope et al. 1996; Zhang and Hewitt 2003). Therefore, the employment of both microsatellites and mtDNA markers allows for a robust understanding of population structure (Hui et al. 2017; Pirog et al. 2019).

The class Ascidiacea (Tunicata, Chordata) is ubiquitous in most marine ecosystems (Shenkar and Swalla 2011), and many of its species play a critical role in ecosystem functioning (Gili and Coma 1998; Loo and Rosenberg 1996; McQuaid and Branch 1985). The Ascidiacea species *Pyura herdmani and Pyura stolonifera* are collectively known as redbait in South Africa. *Pyura herdmani* is also present in northwest Africa (Lafargue and Wahl 1986; Monniot and Bitar 1983; Rius and Teske 2011, 2013; Teske et al. 2011). These filter-feeding sessile species live in intertidal and subtidal habitats wherethey often form vast aggregations of densely spaced individuals (Castilla et al. 2004; Monteiro et al. 2002). They are among the most important bioengineers on rocky shores along southern and north-western African coastlines (Teske et al. 2011). Being a dominant component of benthic communities, they act as bioengineers (Lawton and Jones 1995), occupying much of the available space (Castilla et al. 2004), increasing architectural complexity (Guiñez and Castilla 2001) and enhancing species richness (Cole and McQuaid 2010). The monitoring of population dynamics of these dominant species is important to assess the status of rocky shore ecosystem, which can be affected by a variety of stressors and diseases (Hanekom et al. 1999) and thus impact ecosystem services (Manríquez et al. 2016). In addition, disturbances such as human harvesting may affect the complexity of community assemblies in intertidal and subtidal areas (Rius et al. 2017).

Here, we describe the development of microsatellites and mitochondrial genetic markers for *P. herdmani* and *P. stolonifera*. The microsatellite data were tested on a large number of samples from South Africa’s south-eastern coastline, whereas the mitochondrial markers were developed using a small number of individuals across their range, including *P. herdmani* from northwest Africa. These genetic resources can be used to characterise genetic diversity, demographic history, and spatial structure of both.

**Material and methods**

***Sample acquisition and DNA extraction***

Samples for microsatellite development were collected in September and November 2016 along the south-eastern coastline of South Africa, with a focus on the Wild Coast. This region is located in a transitional zone between the subtropical and warm-temperate coastal bioregions (Emanuel et al. 1992; Turpie et al. 2000) and may represent a distinct bioregion rather than merely an area of faunal overlap (Golla et al. 2020; Jooste et al. 2018). In addition, the Wild coast was not well represented in previous studies of *Pyura* spp. (e.g. Teske et al. 2011). The study area has seen effects of contemporary climate change (Popova et al. 2016), where marine species are adapting in order to survive (Hawes et al. 2018; Miller et al. 2018).

Samples were collected at 13 different sites and identified to species level (Table 1) following (Rius and Teske 2011). Mantle tissue samples of *Pyura* spp*.* were stored in absolute ethanol. DNA was extracted using the CTAB protocol (Doyle and Doyle 1987), and purified using the QIAquick PCR Purification Kit, following the manufacturer’s instructions (Qiagen, Hilden, Germany). The purity and concentration of the DNA was evaluated by measurement of the 260/280 ratio using a NanoDrop®1000 Spectrophotometer 3.7.

*Microsatellite detection*

Four samples of genomic DNA, two for each species of *Pyura*, were sequenced at GenoScreen, Lille, France (www.genoscreen.fr). The development of a microsatellite library required 1 μg of DNA from each sample, and 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries was performed as described in Malausa et al. (2011). In short, total DNA was mechanically fragmented and enriched with AG, AC, AAC, AAG, AGG, ACG, ACAT and ATCT repeat motifs. Thereafter, enriched fragments were amplified, and PCR products were purified and quantified using the QIAquick PCR Purification Kit.

Reads containing microsatellites were analysed using the QDD bioinformatics software (Meglécz et al. 2009). QDD performs all bioinformatics processes from raw sequences to acquiring PCR primers: removal of adapters/vectors, identification of microsatellites, redundancy detection/possible association of mobile components, collection of target microsatellite sequences and first design using BLAST (Altschul et al. 1990), ClustalW (Thompson et al. 2003) and Primer3 (Hancock et al. 2004).

We detected a group of 32 microsatellites, which had a minimum number of seven repeats. We then tested the microsatellite primers using a three-primer amplification protocol (Schuelke 2000), where the forward primer is modified by adding an M13 sequence (TGTAAAACGACGGCCAGT) and its role is taken over by a fluorescently labelled M13 primer once it is used up. In addition, a GTTT PIG-tail (Brownstein et al. 1996) was added to the reverse primer.Each 10 µl of PCR reaction contained 0.1 µM of deoxyribonucleotide triphosphates (dNTPs), 1 µM of 10 x PCR buffer solution (NH4), 1.9 mM of MgCl2, 0.24 mg ml−1 of bovine serum albumin (BSA), 0.13 µM of forward primer, 0.2 µM of reverse primer, 0.06 µM of fluorescently labelled (NED/VIC/FAM/PET) M13 primer, 0.04 U of Super Therm Taq polymerase (Separation Scientific, South Africa) and 5 ng µl−1 of template DNA. The PCR profile consisted of initial denaturation (94 °C for 3 min), 35 cycles of denaturing (94 °C for 45 sec), annealing (1 min at a primer-specific annealing temperature, *Ta*; Table 2) and extension (72 °C for 45 sec), followed by a final extension step (72 °C for 10 min). PCR products were genotyped by the Central Analytical Facility (CAF) at Stellenbosch University, South Africa.

Two hundred and sixty individual samples of *Pyura* spp. were genotyped for the characterisation of the five microsatellite makers that amplified reliably in all samples. These were characterised by perfect repeat patterns (De Barba et al. 2017; Zane et al. 2002), and three of them were trinucleotides, which can be scored more reliably than dinucleotides. Alleles were scored using Geneious R11 (Kearse et al. 2012), and individual genotypes were manually checked and edited. A subset of samples (10%) were genotyped independently by another researcher using a different software (GeneMapper 5, Applied Biosystems) to check for genotyping errors. MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) was used to check for the presence of stutter, null alleles and large allele dropout. The number of alleles (*N*A), observed heterozygosity (*H*O), expected heterozygosity (*H*E) and tests for deviation from Hardy-Weinberg equilibrium (HWE) were conducted in Genalex 6.5 (Peakall and Smouse 2012). Genetix 4.05.2 (Belkhir et al. 2004) was used to calculate the inbreeding coefficient *F*IS with 95% confidence intervals generated over 1 000 bootstrap replicates. GENEPOP 4.7 (Rousset 2008) was used to test for linkage disequilibrium (LD) between pairs of loci with 1000 dememorisations, 500 batches and 10 000 interactions per batch (investigating per individual site and per study population). In order to adjust the *p*-values of tests for HWE and LD for multiple comparisons, Bonferroni correction (Rice 2006) was applied.

*Partial mitogenome assembly*

Ten extracted DNA samples, eight from South Africa and two from Morocco, were used to create libraries using ezRAD (Toonen et al. 2013), and were sequenced at the Genome Centre at the Max Planck Institute for Plant Breeding Research (Cologne, Germany), on an Illumina Hiseq3000 platform (Illumina, USA) using 150 base pair (bp) paired-end chemistry. The ezRAD method does not amplify the whole mitogenome but only the regions flanked by two restriction enzymes (Toonen et al. 2013). Raw reads were analysed with FastQC (Cock et al. 2009), and where removed when the following conditions were met: ≥10% unidentified nucleotides, >50% of bases having phred quality <20, and reads with >10 nt aligned to the adapter (allowing 10% mismatches). Presumed duplicates generated by PCR amplification in the library construction process were also discarded. TRIMMOMATIC (Bolger et al. 2014) was used to remove adapters, and reads shorter than 150bp were discarded. The CLC GENOMIC WORKBENCH 12.0 (CLC Inc., Aarhus, Denmark) with default settings was used for quality control and read assembly. All the resulting reads were mapped against the complete mitochondrial genome of *P. gangelion* (GenBank accession number NC\_021465; Rubinstein et al. 2013). In order to identify the most variable regions, partial mitogenomes were aligned using MUSCLE (Edgar 2004) in MEGA 7 (Kumar et al. 2016). A Bayesian phylogenetic tree was constructed in BEAST 2.5.0 (Bouckaert et al. 2014) using the HKY substitution model (Hasegawa et al. 1984) and 10 million Markov chain Monte Carlo (MCMC) iterations with 20% burn-in. Following an earlier study of ascidian mitogenomes (Griggio et al. 2014), only the concatenated sequences of 13 partial protein-coding genes (PCGs) were used in this analyses, because the large number of missing and ambiguous sites in the rRNAs and tRNAs resulted in these markers being difficult to align. The tree was visualised in Figtree 1.4.3 (Rambaut and Drummond 2019).

**Results**

*Microsatellites*

A total of 4,339 putative microsatellites were detected. Of the 32 primers that were tested, five that amplified reliably in both species and showed polymorphism and heterozygosity features were chosen for detailed analyses. The characteristics of these microsatellites are presented in Table 2. No markers displayed stuttering, null alleles or large allele dropout. A total of 61 alleles were detected across all markers, with an average number of alleles per locus of 12.2. Loci Pstol2 and Pherd2had the lowest number of alleles (6 and 11, respectively), and locus Pherd3the highest (17). *H*O ranged from 0.120 in Pherd2 to 0.561 in Pherd3, and *H*E ranged from 0.190 inPherd2 to 0.757 in Pherd3. Locus Pstol1showed evidence for outbreeding in both populations of *P. herdmani* and a significant inbreeding coefficient (*F*IS) was found at locusPstol2, while *P. stolonifera* showed significant outbreeding at locus Pherd3 and evidence of inbreeding at locus Pherd2. Genotype frequencies at each locus conformed to expectations under HWE, except for Pherd3 (*p* = 0.003) in *P. stolonifera* and Pherd2in *P. herdmani* from the subtropical bioregion (*p* = 0.0002) (Table 3). No significant linkage disequilibrium was detected between the locus pairs, except for loci Pstol1 and Pherd2, but this was only foundin samples of *P. herdmani* from the subtropical bioregion.

*Mitogenome*

The sequencing resulted in a total of 49,327,633 paired-end short reads with an average sequence length of 150 bp and a GC content of 45%. The length of the partial mapped mitogenome for the concatenated sequences of 13 protein-coding genes (PCGs) of *P. stolonifera* and four populations of *P. herdmani* are shown in Table 4 (GenBank accession numbers MT840149-MT840183). Comparing the partial reconstruction to *P. gangelion*, *P. stolonifera* recovered 47.2% of the total sequence, the *P. herdmani* population defined as common temperate 32.6%, rare temperate *P. herdmani* 28.1%, subtropical *P. herdmani* 36.9%, and Moroccan *P. herdmani* 49.9%.

The number of single nucleotide polymorphisms (SNPs) per gene region, and the relative variation when taking the length of each gene fragment into account, are listed in Table 5. Among all protein-coding genes, 2,042 SNPs were detected, of which 757 were parsimony informative. Cox1 was the most variable gene in terms of total number of SNPs (404), and was followed by Nad5 with 288 SNPs, whereas Atp8 had the lowest variability, with only 35 SNPs. However, when length of the gene fragment was taken into account, Atp8 was the most variable gene (35%), followed by Cox2 (27%) and Cox1 (26%).

The phylogenetic tree shown in Figure 1 was inferred using the concatenated partial sequences of the 13 PCGs. It provided a well-resolved phylogeny, with all nodes being strongly supported (posterior probability [PP] > 0.99). *Pyura gangelion* was used as the outgroup species and *P. stolonifera* and *P. herdmani* formed two distinct monophyletic groups, and *P. herdmani* is further subdivided into distinct populations.

**Discussion**

In this study, we tested new genetic markers with South African and northwest African individuals of *P. stolonifera* and *P. herdmani.* The microsatellites could be used to detect fine-scale population structure and very recent evolutionary changes (Vieira et al. 2016; Giles et al. 2018) in order to gain a better understanding of the status of the marine ecosystem in the context of climate change, as ~~was~~ has been done for other species (Tavares et al. 2018). In the study area in particular, much of which is located in a biogeographical transition zone between the subtropical and warm-temperate coastal bioregions, and contemporary climate change has increased the sea temperature and intensified the warm, southward-flowing Agulhas Current (Popova et al. 2016). The studied species have a relatively narrow temperature tolerance range (Hudson et al. 2021), suggesting high vulnerability to warming. The partial reconstruction of the mtDNA from NGS data allowed a comparison of genetic variation contained within different mtDNA markers. While ezRAD and similar methods are not suitable for the reconstruction of complete mitogenomes (Shafer et al. 2017), they are a useful tool in phylogeny reconstruction (Cariou et al. 2013). Utilising new software and pipeline procedures (Toonen et al. 2013; Terraneo et al. 2018; Stobie et al. 2019), we were able to reconstruct partial sequences of 13 PCGs of the five study populations of *Pyura* spp. Despite missing data**,** the partial PCG sequences were suitable to compare variability between different genes (Rubinstein et al. 2013; Ip et al. 2019). Correcting for sequence length, we found the highest numbers of SNPs in the Atp8, Cox2 and Cox1 regions, rendering those genes the most informative markers for more detailed phylogeographical analyses. The development of the phylogenetic tree with the 13 PDGs (Rubinstein et al. 2013) of the *Pyura* spp. further helped to determine evolutionary relationships between the study populations. Our findings show how different populations of *P. herdmani* are rather divergent, indicating the need for further studies to determine whether they could be reproductively isolated, distinct species. The genetic tools developed here can be used in future studies on *Pyura* spp. to locate and design additional primers to specifically target the most variable mitochondrial regions, as was indicated in a similar study by Klein et al. (2019). These new genomic resources, combined with previous nuclear and mitochondrial markers, will facilitate the fine-scale assessment of population structure and analyses of the evolution of *Pyura* spp. on a regional and global scale.

In conclusion, considering that *Pyura* spp. are essential bioengineers ~~for~~ on the South African coastline (Rius et al. 2017), these tools ~~will also~~ allow a better understanding of how the ranges of the morphologically very similar but genetically distinct populations of P. herdmani may shift under conditions of climate change. We believe that monitoring ~~of~~ these key species ~~can~~ could highlight their role in the coastal and marine ecosystem, which could contribute to achieving ~~an~~ the overarching goal of the preservation of ~~the~~ ecosystem functioning ~~of~~ on the South African coastline.

*Acknowledgements* — This work was supported by the National Research Foundation of South Africa (Competitive Support for Rated Researchers Grant No. 98893 to MR and PRT) and the University of Johannesburg (FRC/URC Grant to PRT). AD is grateful to the University of Johannesburg for awarding her a Global Excellence and Stature (GES) PhD bursary. We thank the CSIR’s Centre for High Performance Computing (CHPC) for computational resources. Special thanks are given to J Klein for helping with the collection of samples, and to A Emami-Khoyi, TR Golla, D Monsanto, BJ van Vuuren, and J Sandoval-Castillo for consultations regarding mitogenome assemblage ~~consultation~~. We also thank the Genome Centre at the Max Planck Institute for Plant Breeding Research in Cologne, Germany, for the Illumina sequencing, and two anonymous reviewers for useful comments.

**ORCID**

Alessia Dinoi: https://orcid.org/0000-0001-8122-3625

Marc Rius: https://orcid.org/0000-0002-2195-6605

Mbaye Tine: https://orcid.org/0000-0003-2427-1027

Peter R Teske: https://orcid.org/0000-0002-2838-7804

**References**

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Avise JC. 1994. Molecular markers, natural history and evolution. Chapter 5. Chapman and Hall, New York, USA. *-*.

Belkhir K, Borsa N, Chikhi L, Raufaste N, Bonhomme F. 2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. *Laboratoire Génome, Populations, Interactions, CNRS UMR 5000*, 19.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Brownstein MJ, Carpten JD, Smith JR. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques*, *20*(6), 1004–1010.

Cariou M, Duret L, Charlat S. 2013. Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. *Ecology and Evolution*, *3*(4), 846–852. https://doi.org/10.1002/ece3.512

Castilla JC, Lagos NA, Cerda M. 2004. Marine ecosystem engineering by the alien ascidian *Pyura praeputialis* on a mid-intertidal rocky shore. *Marine Ecology Progress Series*, *268*, 119–130.

Cock PJA, Fields CJ, Goto N, Heuer ML, Rice PM. 2009. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, *38*(6), 1767–1771. https://doi.org/10.1093/nar/gkp1137

Cole VJ, McQuaid CD. 2010. Bioengineers and their associated fauna respond differently to the effects of biogeography and upwelling. *Ecology*, *91*(12), 3549–3562. https://doi.org/10.1890/09-2152.1

De Barba M, Miquel C, Lobréaux S, Quenette PY, Swenson JE, Taberlet P. 2017. High-throughput microsatellite genotyping in ecology: improved accuracy, efficiency, standardization and success with low-quantity and degraded DNA. *Molecular Ecology Resources*, *17*(3), 492–507. https://doi.org/10.1111/1755-0998.12594

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, *19*, 11–15.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Emanuel BP, Bustamante RH, Branch GM, Eekhout S, Odendaal FJ. 1992. A zoogeographic and functional approach to the selection of marine reserves on the west coast of South Africa. *South African Journal of Marine Science*, *12*(1), 341–354. https://doi.org/10.2989/02577619209504710

Giles EC, Petersen-Zúñiga C, Morales-González S, Quesada-Calderón S, Saenz-Agudelo P. 2018. Novel microsatellite markers for *Pyura chilensis* reveal fine-scale genetic structure along the southern coast of Chile. *Marine Biodiversity*, *48*(4), 1777–1786. https://doi.org/10.1007/s12526-017-0672-9

Gili JM, Coma R. 1998. Benthic suspension feeders: Their paramount role in littoral marine food webs. *Trends in Ecology and Evolution*, Vol. 13, pp. 316–321. https://doi.org/10.1016/S0169-5347(98)01365-2

Golla TR, Pieterse L, Jooste CM, Teske PR. 2020. Discovery of populations endemic to a marine biogeographical transition zone. *Diversity and Distributions*, *26*(12), 1825–1832. https://doi.org/10.1111/ddi.13162

Goodwin KD, Thompson LR, Duarte B, Kahlke T, Thompson AR, Marques JC, Caçador I. 2017. DNA sequencing as a tool to monitor marine ecological status. *Frontiers in Marine Science*, Vol. 4, p. 107. https://doi.org/10.3389/fmars.2017.00107

Griggio F, Voskoboynik A, Iannelli F, Justy F, Tilak M-K, Xavier T, Pesole G, Douzery EJP, Mastrototaro F, Gissi C. 2014. Ascidian mitogenomics: comparison of evolutionary rates in closely related taxa provides evidence of ongoing speciation events. *Genome Biology and Evolution*, *6*(3), 591–605. https://doi.org/10.1093/gbe/evu041

Guiñez R, Castilla JC. 2001. An allometric tridimensional model of self-thinning for a gregarious tunicate. *Ecology*, *82*(8), 2331–2341. https://doi.org/10.1890/0012-9658(2001)082[2331:AATMOS]2.0.CO;2

Hancock JM, Zvelebil MJ, Hancock JM. 2004. PRIMER3. In *Dictionary of Bioinformatics and Computational Biology*. https://doi.org/10.1002/9780471650126.dob0560.pub2

Hanekom N, Harris JM, Branch GM, Allen JC. 1999. Mass mortality and recolonization of *Pyura stolonifera* (Heller) on the South Coast of South Africa. *South African Journal of Marine Science*, (21), 117–133. https://doi.org/10.2989/025776199784125953

Hasegawa M, Kishino H, Yano T. 1984. Dating of the Human-Ape Splitting by a Molecular Clock. *Japjgen*, *59*(6), 616.

Hawes NA, Tremblay LA, Pochon X, Dunphy B, Fidler AE, Smith KF. 2018. Effects of temperature and salinity stress on DNA methylation in a highly invasive marine invertebrate, the colonial ascidian *Didemnum vexillum*. *PeerJ*, *6*, e5003. https://doi.org/10.7717/peerj.5003

Hudson J, McQuaid CD, Rius M (2021) Contemporary climate change hinders hybrid performance of ecologically dominant marine invertebrates. Journal of Evolutionary Biology, 34, 60-72

Hui M, Nuryanto A, Kochzius M. 2017. Concordance of microsatellite and mitochondrial DNA markers in detecting genetic population structure in the boring giant clam Tridacna crocea across the Indo-Malay Archipelago. *Marine Ecology*, *38*(1), e12389. https://doi.org/10.1111/maec.12389

Ip YCA, Tay YC, Gan SX, Ang HP, Tun K, Chou LM, Huang D, Meier R. 2019. From marine park to future genomic observatory? Enhancing marine biodiversity assessments using a biocode approach. *Biodiversity Data Journal*, *7*, e46833–e46833. https://doi.org/10.3897/BDJ.7.E46833

Jooste CM, Oliver J, Emami-Khoyi A, Teske PR. 2018. Is the Wild Coast in eastern South Africa a distinct marine bioregion? *Helgoland Marine Research*, *72*(1). https://doi.org/10.1186/s10152-018-0509-3

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Klein JD, Bester-van der Merwe AE, Dicken ML, Emami-Khoyi A, Mmonwa KL, Teske PR. 2019. Genomic resources for the spotted ragged-tooth shark *Carcharias taurus*. *African Journal of Marine Science*, *41*(1), 115–118. https://doi.org/10.2989/1814232X.2019.1576541

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

Lafargue F, Wahl M. 1986. Contribution to the knowledge of littoral ascidians (Ascidiacea, Tunicata) of the Senegalese coast. *Bulletin de l’Institut Français d’Afrique Noire*, *46A*(3–4), 385–402.

Lawton JH, Jones CG. 1995. Linking Species and Ecosystems: Organisms as Ecosystem Engineers. In Jones Clive G. & J. H. and Lawton (Eds.), *Linking Species & Ecosystems* (pp. 141–150). https://doi.org/10.1007/978-1-4615-1773-3\_14

Loo L-O, Rosenberg R. 1996. Production and energy budget in marine suspension feeding populations: *Mytilus edulis*, *Cerastoderma edule*, *Mya arenaria* and *Amphiura filiformis*. *Journal of Sea Research*, *35*(1–3), 199–207. https://doi.org/10.1016/S1385-1101(96)90747-9

Manríquez PH, Castilla JC, Ortiz V, Jara ME. 2016. Empirical evidence for large-scale human impact on intertidal aggregations, larval supply and recruitment of *Pyura praeputialis* around the Bay of Antofagasta, Chile. *Austral Ecology*, *41*(6), 701–714. https://doi.org/10.1111/aec.12359

McQuaid C, Branch G. 1985. Trophic structure of rocky intertidal communities response to wave action and implications for energy flow. *Marine Ecology Progress Series*, *22*, 153–161. https://doi.org/10.3354/meps022153

Meglécz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin JF. 2009. QDD: A user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics*, *26*(3), 403–404. https://doi.org/10.1093/bioinformatics/btp670

Miller DD, Ota Y, Sumaila UR, Cisneros-Montemayor AM, Cheung WWL. 2018. Adaptation strategies to climate change in marine systems. *Global Change Biology*, *24*(1), e1–e14. https://doi.org/10.1111/gcb.13829

Monniot C, Bitar G. 1983. Sur la presence de Pyura stolonifera (Tunicata, Ascidiacea) a ras achaccar (cot Nord Atlantique Marocaine). Comparaison anatomique distinctive avec Pyura praeputialis. *Bulletin de l’Institut Scientifique,* *7*(Rabat), 83–91. https://doi.org/10.1007/s13398-014-0173-7.2

Monteiro SM, Chapman MG, Underwood AJ. 2002. Patches of the ascidian Pyura stolonifera (Heller, 1878): Structure of habitat and associated intertidal assemblages. *Journal of Experimental Marine Biology and Ecology*, *270*(2), 171–189. https://doi.org/10.1016/S0022-0981(02)00020-5

Peakall R, Smouse PE. 2012. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pirog A, Ravigné V, Fontaine MC, Rieux A, Gilabert A, Cliff G, Clua E, Daly R, Heithaus MR, Kiszka JJ, Matich P, Nevill JEG, Smoothey AF, Temple AJ, Berggren P, Jaquemet S, Magalon H. 2019. Population structure, connectivity, and demographic history of an apex marine predator, the bull shark Carcharhinus leucas. *Ecology and Evolution*, *9*(23), 12980–13000. https://doi.org/10.1002/ece3.5597

Pope LC, Sharp A, Moritz C. 1996. Population structure of the yellow-footed rock-wallaby *Petrogale xanthopus* (Gray, 1854) inferred from mtDNA sequences and microsatellite loci. *Molecular Ecology*, *5*(5), 629–640. https://doi.org/10.1111/j.1365-294X.1996.tb00358.x

Popova E, Yool A, Byfield V, Cochrane K, Coward AC, Salim SS, Gasalla MA, Henson SA, Hobday AJ, Pecl GT, Sauer WH, Roberts MJ. 2016. From global to regional and back again: Common climate stressors of marine ecosystems relevant for adaptation across five ocean warming hotspots. *Global Change Biology*, *22*(6), 2038–2053. https://doi.org/10.1111/gcb.13247

Rambaut A, Drummond AJ. 2019. *FigTree v1. 4.3. 2016*.

Rice WR. 2006. Analyzing Tables of Statistical Tests. *Evolution*, *43*(1), 223. https://doi.org/10.2307/2409177

Rius M, Teske PR. 2011. A revision of the *Pyura stolonifera* species complex (Tunicata, Ascidiacea), with a description of a new species from Australia. *Zootaxa*, *2754*(2754), 27–40.

Rius M, Teske PR. 2013. Cryptic diversity in coastal Australasia: a morphological and mitonuclear genetic analysis of habitat-forming sibling species. *Zoological Journal of the Linnean Society*, *168*(3), 597–611. https://doi.org/10.1111/zoj.12036

Rius M, Teske PR, Manríquez PH, Suárez-Jiménez R, McQuaid CD, Castilla JC. 2017. Ecological dominance along rocky shores, with a focus on intertidal ascidians. In *Oceanography and Marine Biology: An Annual Review*. https://doi.org/10.1201/b21944

Rousset F. 2008. GENEPOP’007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Rubinstein ND, Feldstein T, Shenkar N, Botero-Castro F, Griggio F, Mastrototaro F, Delsuc F, Douzery EJP, Gissi C, Huchon D. 2013. Deep sequencing of mixed total DNA without barcodes allows efficient assembly of highly plastic Ascidian mitochondrial genomes. *Genome Biology and Evolution*, *5*(6), 1185–1199. https://doi.org/10.1093/gbe/evt081

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments - A poor man’ s approach to genotyping for research and high-throughput diagnostics. *Nature Biotechnology*, *18*(18), 233–234. Retrieved from http://dx.doi.org/10.1038/72708

Shafer ABA, Peart CR, Tusso S, Maayan I, Brelsford A, Wheat CW, Wolf JBW. 2017. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, *8*(8), 907–917. https://doi.org/10.1111/2041-210X.12700

Shenkar N, Swalla BJ. 2011. Global diversity of Ascidiacea. *PLoS ONE*, 2011/06/20, Vol. 6, pp. e20657–e20657. https://doi.org/10.1371/journal.pone.0020657

Stobie CS, Cunningham MJ, Oosthuizen CJ, Bloomer P. 2019. Finding stories in noise: Mitochondrial portraits from RAD data. *Molecular Ecology Resources*, *19*(1), 191–205. https://doi.org/10.1111/1755-0998.12953

Tavares AI, Nicastro KR, Kolzenburg R, Ragazzola F, Jacinto R, Zardi GI. 2018. Isolation and characterization of nine microsatellite markers for the red alga *Corallina officinalis*. *Molecular Biology Reports*, *45*(6), 2791–2794. https://doi.org/10.1007/s11033-018-4353-y

Terraneo TI, Arrigoni R, Benzoni F, Forsman ZH, Berumen ML. 2018. Using ezRAD to reconstruct the complete mitochondrial genome of *Porites fontanesii* (Cnidaria: Scleractinia). *Mitochondrial DNA Part B: Resources*, *3*(1), 173–174. https://doi.org/10.1080/23802359.2018.1437805

Teske PR, Beheregaray LB. 2009. Intron-spanning primers for the amplification of the nuclear ANT gene in decapod crustaceans. *Molecular Ecology Resources*, *9*(3), 774–776. https://doi.org/10.1111/j.1755-0998.2009.02534.x

Teske PR, Rius M, McQuaid CD, Styan CA, Piggott MP, Benhissoune S, Fuentes-Grünewald C, Walls K, Page M, Attard CRM, Cooke GM, McClusky CF, Banks SC, Barker NP, Beheregaray LB. 2011. “Nested” cryptic diversity in a widespread marine ecosystem engineer: a challenge for detecting biological invasions. *BMC Evolutionary Biology*, *11*(1), 176. https://doi.org/10.1186/1471-2148-11-176

Thompson JD, Gibson TJ, Higgins DG. 2003. Multiple Sequence Alignment Using ClustalW and ClustalX. *Current Protocols in Bioinformatics*, *00*(1), 2.3.1-2.3.22. https://doi.org/10.1002/0471250953.bi0203s00

Toonen RJ, Puritz JB, Forsman ZH, Whitney JL, Fernandez-Silva I, Andrews KR, Bird CE. 2013. ezRAD: A simplified method for genomic genotyping in non-model organisms. *PeerJ*, *2013*(1). https://doi.org/10.7717/peerj.203

Turpie JK, Beckley LE, Katua SM. 2000. Biogeography and the selection of priority areas for conservation of South African coastal fishes. *Biological Conservation*, *92*(1), 59–72. https://doi.org/10.1016/S0006-3207(99)00063-4

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3), 535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Vieira MLC, Santini L, Diniz AL, Munhoz C de F. 2016, July 1. Microsatellite markers: What they mean and why they are so useful. *Genetics and Molecular Biology*, Vol. 39, pp. 312–328. https://doi.org/10.1590/1678-4685-GMB-2016-0027

Weber JL, Wong C. 1993. Mutation of human short tandem repeats. *Human Molecular Genetics*, *2*(8), 1123–1128. https://doi.org/10.1093/hmg/2.8.1123

Zane L, Bargelloni L, Patarnello T. 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology*, *11*(1), 1–16. https://doi.org/10.1046/j.0962-1083.2001.01418.x

Zhang DX, Hewitt GM. 2003, March 1. Nuclear DNA analyses in genetic studies of populations: Practice, problems and prospects. *Molecular Ecology*, Vol. 12, pp. 563–584. https://doi.org/10.1046/j.1365-294X.2003.01773.x

**Figure legend**

**Figure 1:** Bayesian phylogenetic tree, based on 13 partial concatenated protein-coding genes (PCGs), of *Pyura stolonifera and P. herdmani* from four different bioregions. The tree was constructed using the HKY substitution model. The Markov chain Monte Carlo method was set with 10 million iterations following a 20% burn-in. *Pyura gangelion* was used as the outgroup. Posterior probability values are shown at each branch node. The scale axis shows the substitution rate

**Table 1:** Sampling sites and their geographical locations at which samples of *Pyura* spp. were collected in September–November 2016. Although the boundaries of the Wild Coast are poorly defined, samples from Mbizana to approximately East London are considered to be located within this biogeographical transition zone

|  |  |  |
| --- | --- | --- |
| Biogeographical entity | Sampling site | Geographic location (Latitude/Longitude) |
| Subtropical Natal Bioregion | Mzumbe Mouth | 30°45'11" S 30°27'7" E |
| Wild Coast transition zone | Mbizana Mouth | 31°6'11" S 30°10'48" E |
| Wild Coast transition zone | Cathedral Rock | 31°26'41" S 29°46'18" E |
| Wild Coast transition zone | Port St Johns | 31°37'48" S 29°33'5" E |
| Wild Coast transition zone | Coffee Bay | 31°59'6" S 29°9'11"E |
| Wild Coast transition zone | Mazzepa Bay | 32°28'54'' S 28°38'41'' E |
| Wild Coast transition zone | Qolora Mouth | 32°39'23.25" S 28°25'0.17" E |
| Wild Coast transition zone | Morgan Bay | 32°42'39'' S 28°20'27" E |
| Wild Coast transition zone | Haga Haga | 32°45'54.20" S 28°14'58.60" E |
| Wild Coast transition zone | East London | 33° 0'57.35" S 27°55'16.40" E |
| Warm-temperate Agulhas Bioregion | Hamburg | 33°17'50.63" S 27°28'49.66" E |
| Warm-temperate Agulhas Bioregion | Great Fish River Mouth | 33°28'56.61" S 27°9'7.25" E |
| Warm-temperate Agulhas Bioregion | Port Elizabeth | 33°59'31.31" S 25°40'42.98" E |

**Table 2:** Characteristics of the five microsatellite markers developed for *Pyura stolonifera* and *P. herdmani. N*A = total number of alleles, *H*O = observed heterozygosity, *H*E = expected heterozygosity, *Ta* = primer-specific annealing temperature

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Locus | Primer sequence | Repeat pattern | Allele size range | *Ta* (°C) | *N*A | *H*O |  | *H*E |
|  |  |  |  |  |  |  |  |  |
| Pstol1 | F-ACAACCATCACCCTGATTGC | (AC)11 | 73–137 | 62 | 15 | 0.363 |  | 0.372 |
|  | R-TCGAGTGAGAAGACGTGGTG |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| Pstol2 | F-GTTGATTTACAGCGAGGGTTG | (TTC)7 | 64–112 | 52 | 6 | 0.165 |  | 0.384 |
|  | R-TTCGAATCAGGAATCCACCT |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| Pherd1 | F-ATCTCCCGCAGATTCTTCCT | (TCT)7 | 136–187 | 62 | 12 | 0.483 |  | 0.647 |
|  | R-TATTTTGGGGTTCAAGCAGC |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| Pherd2 | F-CGGATAATCTCACTCGTGGAA | (AG)17 | 92–120 | 52 | 11 | 0.120 |  | 0.190 |
|  | R-GAGAATAGAATAGAGCACAATAACAA |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| Pherd3 | F-GAGTTGACTGAATTCCCATCG | (TAG)12 | 93–171 | 60 | 17 | 0.561 |  | 0.757 |
|  | R-TCACAGTTGCTTCATATTCGC |  |  |  |  |  |  |  |

**Table 3:** Hardy-Weinberg equilibrium (HWE) and inbreeding coefficient(*F*IS) of the five microsatellite markers developed for *Pyura stolonifera* andthe two populations of *P. herdmani*; *p*HWE = *p-*value for Hardy-Weinberg equilibrium, \*significant after Bonferroni correction at *p* < 0.005.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Locus | *F*IS (95% CI)  *P. stolonifera* | *P. herdmani* (temperate) | *P. herdmani subtropical* | *p*HWE  *P. stolonifera* | *P. herdmani* (temperate) | *P. herdmani (*subtropical) |
| Pstol1 | 0.285 (0.018, 0.523) | –0.124 (–0.215, –0.056) | –0.055 (–0.320, 0.155) | 0.426 | 0.437 | 0.238 |
| Pstol2 | –0.043 (–0.312, 0.000) | 0.839 (–0.043, 1.000) | 0.848 (0.000, 1.000) | 0.261 | 0.061 | 0.047 |
| Pherd1 | 0.130 (–0.176, 0.408) | 0.471 (0.315, 0.636) | 0.292 (–0.008, 0.545) | 0.407 | 0.268 | 0.444 |
| Pherd2 | 0.654 (–0.027, 1.000) | 0.000 (–0.037, 0.000) | 0.541 (0.228, 0.774) | 0.429 | 0.145 | <0.001\* |
| Pherd3 | –0.529 (–1.000, –0.047) | 0.512 (0.298, 0.685) | 0.060 (–0.125, 0.220) | 0.003\* | 0.191 | 0.085 |

**Table 4:** Partial mitogenome length for the concatenated 13 protein-coding genes (PCGs) of *Pyura stolonifera* and *P. herdmani* from four different bioregions

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Species | Population defined by Teske et al. 2011 | Bioregion | Geographical origin | Site name | Mitogenome length (bp) |
| *P. stolonifera* | – | Cool temperate | Western South Africa | False Bay | 5 141 |
| *P. herdmani* | Temperate common | Warm temperate | Southern South Africa | Knysna | 3 549 |
| *P. herdmani* | Temperate rare | Cool temperate | Western South Africa | Simon's Town | 3 065 |
| *P. herdmani* | Subtropical | Subtropical | Southeastern South Africa | Morgan Bay | 4 016 |
| *P. herdmani* | Moroccan | Moroccan | Northwestern Africa | La Madrague | 5 432 |

**Table 5:** The number of single nucleotide polymorphisms (SNPs) for each of the 13 protein-coding genes (PCGs) and the relative variation of the five partial mitogenomes mapped for *Pyura* spp. Atp = ATP synthase, Cob = Cytochrome oxidase b, Cox = Cytochrome c oxidase, Nad = NADH dehydrogenase

|  |  |  |  |
| --- | --- | --- | --- |
| Gene region | Average length (bp) | Total SNPs | Relative variation (%) |
| Atp6 | 630 | 86 | 13.65 |
| Atp8 | 100 | 35 | 35.00 |
| Cob | 1104 | 275 | 24.90 |
| Cox1 | 1551 | 404 | 26.04 |
| Cox2 | 678 | 180 | 26.54 |
| Cox3 | 789 | 161 | 20.40 |
| Nad1 | 915 | 182 | 19.89 |
| Nad2 | 999 | 132 | 13.21 |
| Nad3 | 390 | 41 | 10.51 |
| Nad4 | 1324 | 138 | 10.42 |
| Nad4l | 234 | 36 | 15.38 |
| Nad5 | 1734 | 288 | 16.60 |
| Nad6 | 435 | 84 | 19.31 |