



## Preliminary population genomic study on the sandfish *Holothuria (Metriatyla) scabra*

Monal M. Lal<sup>\*,†</sup> , Deo A. S. Macahig<sup>‡</sup>, Kelly T. Brown<sup>†</sup> , Marie A. Juinio-Meñez<sup>‡</sup> , Paul C. Southgate<sup>\*</sup> and Rachel Ravago-Gotanco<sup>‡</sup> 

<sup>\*</sup>Australian Centre for Pacific Islands Research and School of Science, Technology and Engineering, University of the Sunshine Coast, 90 Sippy Downs Drive, Maroochydore, QLD 4558, Australia; <sup>†</sup>Discipline of Marine Studies, School of Agriculture, Geography, Environment, Ocean and Natural Sciences, University of the South Pacific, Lower Laucala Campus, Laucala Bay Road, Suva, Fiji Islands; <sup>‡</sup>The Marine Science Institute, University of the Philippines Diliman, Quezon City, Philippines

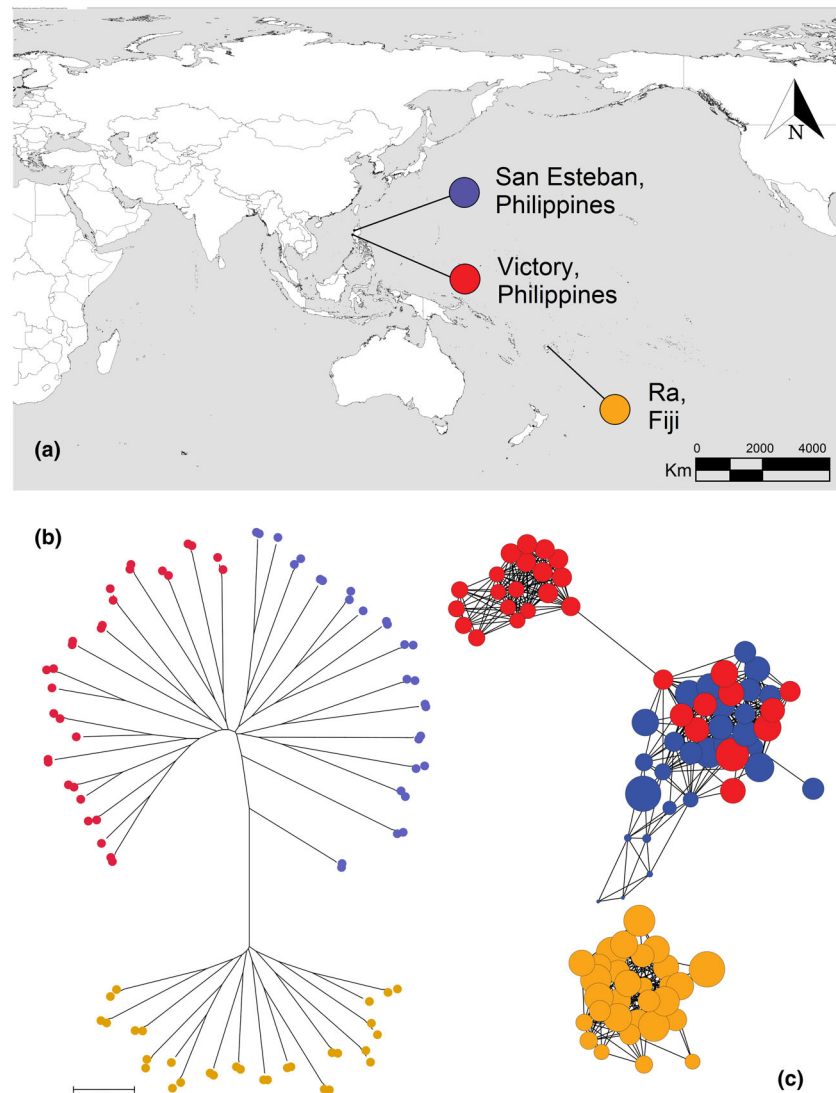
Accepted for publication 25 June 2021

Population genetic data are vital in fisheries stock assessments, as they offer unique insights into stock health and

structure by elucidating levels of genetic diversity, gene flow and the presence of local adaptation.<sup>1,2</sup> Many marine invertebrates, such as the sea cucumber *Holothuria (Metriatyla) scabra* or sandfish, have remained non-model organisms for which genetic information is unavailable.<sup>3–6</sup> Sandfish and other sea cucumbers are highly valued in the global bêche-de-mer trade and are consequently heavily fished.<sup>7,8</sup> Recently established cultured populations may alleviate fishing pressure on wild stocks and replenish extirpated populations.

We undertook a preliminary assessment of genetic diversity, structure and relatedness in the sandfish. DNA was extracted using the Qiagen DNEasy Blood and Tissue kit and CTAB chloroform–isoamyl alcohol protocol with modifications.<sup>9</sup> We used DARtseq<sup>10</sup> in order to assess genetic diversity, neutral and adaptive structure and relatedness in one hatchery-produced (Philippine) and two wild (Fiji and the Philippines) populations (Appendix S1).

**Figure 1** (a) Sample collection locations for individuals of *Holothuria scabra* genotyped. Colour codes represent San Esteban ( $n = 28$ ) in blue and Victory Sea Ranch, Bolinao in red ( $n = 30$ ), both in the Philippines. Ra, Viti Levu, Fiji ( $n = 30$ ) is represented in gold. (b) Neighbour-joining tree visualising genetic relationships among individuals at 164  $F_{ST}$  outlier SNPs putatively under adaptive selection. The scale bar denotes tree branch lengths which represent one minus the proportion of shared alleles (1-PSA) genetic distance. Genetic distances were computed using the R package ADEGENET and the tree was generated in MEGA6. (c) Fine-scale Netview R population genetic structure network of *H. scabra* individuals drawn at mutual- $k$  nearest neighbour (mk-NN) threshold = 19 using 8725 selectively neutral SNPs. Node sizes have been mapped to the relatedness of individual animals (neighbourhood connectivity based on average proportion of shared alleles identical by state - IBS distances), and the network constructed using the organic topology framework option in CYTOSCAPE v.2.8.3.



A total of 90 921 SNPs were identified and genotyped across 88 individuals. After filtering SNPs at 95% call rate, the average call rate of individuals was 71% with an average polymorphic information content of 0.156, indicating that the DArTseq platform is robust for generating large numbers of informative SNPs (Fig. S2a,b). Following the final filtering steps (read depth >8, average polymorphic information content >1%, MAF >2% per population and average repeatability >95%), 8725 selectively neutral and 164  $F_{ST}$  outlier SNPs remained for further analysis. Substantial differentiation (average pairwise  $F_{ST}$  = 0.325; Table S2, Fig. 1c) was detected between the two countries. In addition, we observed adaptive variation (164  $F_{ST}$  outlier SNPs; Fig. 1b). Spatial patterns of diversity are consistent with geographic separation between sites. The relatively close relationship of the wild and captive Philippines populations is supported by a previous microsatellite-based study.<sup>1</sup> We did not observe a decrease in heterozygosity of the captive population (Table S1). However, the captive effective population size is much lower than that in the wild ( $N_{eLD}$  from about 2000 to 4000 compared with ~19), suggesting that broodstock selection and hatchery practices should be carefully managed.

We conclude that high-resolution population genomic analysis using DArTseq holds promise for application to other marine invertebrates, and particularly non-model taxa. This information is imperative for defining conservation and management policy, addressing stock recovery measures and developing guidelines for responsible aquaculture.

**Funding:** The Australian Centre for International Agricultural Research (ACIAR) project FIS/2016/122: 'Increasing technical skills supporting community-based sea cucumber production in Vietnam and the Philippines' supported this study. The funding source (ACIAR) had no direct involvement in the study design, collection, analysis and interpretation of the data, nor the decision to submit this article for publication.

#### Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Ethics approval

All sandfish were handled in accordance with the University of the Sunshine Coast's animal ethics requirements and guidelines, in compliance with the Australian Code for the Care and Use of Animals for Scientific Purposes 8th Edition (2013).

#### Data availability statement

Genotypic data for sandfish sampled and sequenced during this study are available as Appendix S2. Filtered SNP datasets generated from sequencing have been provided as Appendices S3 and S4 for selectively neutral and putatively adaptive markers, respectively.

#### References

- 1 Ravago-Gotanco R. & Kim K.M. (2019) *Fish Res* **209**, 143–55.
- 2 Lal M.M. *et al.* (2016) *Mar Genomics* **25**, 57–68.
- 3 Mishra A. *et al.* (2019) *Indian J Anim Sci* **89**, 699–701.
- 4 Schiebelhut L.M. *et al.* (2017) *Mol Ecol Resour* **17**, 721–9.
- 5 Madeira P. *et al.* (2018) *Eur Zool J* **85**, 89–94.
- 6 Nowland S.J. *et al.* (2015) *Invertebr Biol* **134**, 252–9.
- 7 Pakoa K. & Bertram I. (2013) *SPC Beche-de-mer Inf Bull* **33**, 49–52.
- 8 Purcell S.W. *et al.* (2013) *Fish Fish* **14**, 34–59.
- 9 Lal M.M. *et al.* (2016) *PLoS ONE* **11**, e0161390.
- 10 Kilian A. *et al.* (2012) *Methods Mol Biol* **888**, 67–89.
- 11 Lind C.E. *et al.* (2017) *Anim Genet* **48**, 362–4.

Correspondence: M. M. Lal (mlal1@usc.edu.au)

#### Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Appendix S1.** Methods, materials and detailed results.

**Appendix S2.** Sequence and SNP data for 88 *Holothuria scabra* individuals sampled. This dataset contains 40 200 genome-wide SNPs supplied by DArTseq genotyping that were filtered to remove duplicate (clone SNPs) and retained at a global call rate of 95% (see Appendix S1 for further details). Columns C and D contain the allele and trimmed sample sequences respectively, with column E containing the SNP call and SNP base pair position with respect to the allele and trimmed sequences. 11 SNP-specific quality metrics are provided from columns G to R, and sample names are given from column S onwards. As per standard DArT PL reporting, genotype scores of '0', '1', '2' and '—' indicate a reference sequence allele homozygote, SNP allele homozygote, heterozygote and missing data, respectively.

**Appendix S3.** Selectively neutral genotypic data. Genotypes of 88 individuals of *H. scabra* at 8725 selectively neutral genome-wide SNPs are provided in a standard STRUCTURE format.

**Appendix S4.** Putatively adaptive genotypic data. Genotypes of 88 individuals of *H. scabra* at 164  $F_{ST}$  outlier genome-wide SNPs are provided in a standard STRUCTURE format.