Identification of SNP markers for population genetics studies of the Mediterranean gorgonian, Leptogorgia sarmentosa

M. Padrón, M. Milhes, A. Delcamp, E. Guichoux

Conservation Genetics Resources
https://doi.org/10.1007/s12686-021-01218-3

TECHNICAL NOTE

Identification of SNP markers for population genetics studies of the Mediterranean gorgonian, *Leptogorgia sarmentosa*

M. Padrón[1] · M. Milhes[2] · A. Delcamp[3] · E. Guichoux[3]

Received: 24 December 2020 / Accepted: 21 May 2021

Abstract
A set of single nucleotide polymorphisms (SNP) was developed from the transcriptome of the gorgonian *Leptogorgia sarmentosa*. 21,491 putative SNPs were identified, and a selected set of 320 SNPs was tested using a MassARRAY System on 95 samples from the NW Mediterranean. A total of 153 SNPs were successfully genotyped and found to be polymorphic. The newly developed loci will be a valuable tool for population genetics studies of one of the most ubiquitous gorgonian species in the Mediterranean Sea.

Keywords *Leptogorgia sarmentosa* · Single Nucleotide Polymorphism (SNP) · MassARRAY · gorgonian

Introduction

Gorgonians are conspicuous components of littoral benthic ecosystems in temperate, tropical and polar areas. They contribute significantly to providing habitats for the epifauna and to increasing the biomass and diversity of the benthic community (True 1970; Kinzie 1973; Weinberg 1978; Starmans et al. 1999). *Leptogorgia sarmentosa* Esper 1789 is a common member of the benthic Mediterranean fauna that is usually found between 20 and 200 m depth. It is the only species of the genus *Leptogorgia* described in the Mediterranean (Carpine 1963; Carpine and Grasshoff 1975; Weinberg 1976). Given the complex systematic history of the genus, most of the molecular studies have been focused on discerning the phylogenetic relationships among its species (Breedy and Guzmán 2007). The mitogenome of *L. sarmentosa* is the only source of molecular data available for the species to date (Poliseno et al. 2017). Here, we sequence the transcriptome of this species, and present a set of SNP markers developed specifically for population genetics studies of *L. sarmentosa*. These results will provide a very useful tool for understanding the genetic variability of the species populations. Moreover, given the transcriptomic origin of the SNP markers, it is also likely that these markers would help analyzing the genetic differentiation of closely related species.

✉ M. Padrón
mariana.padron@gmail.com

[1] Laboratoire d'Ecogeochimie des Environnements Benthique, CNRS, Sorbonne Université, UMR 8222, 66650 Banyuls-sur-Mer, France

[2] IINRA, GeT-PlaGe, Genotoul, 31326 Castanet-Tolosan, France

[3] Univ. Bordeaux, INRAE, BIOGECO, 33610 Cestas, France

Materials and methods

Transcriptome sequencing

The transcriptome of *L. sarmentosa* was sequenced from the total RNA of ten individuals collected in Banyuls sur Mer, France. Small colonies were collected by SCUBA diving between 10 and 20 meters' depth. The samples were brought alive to the laboratory. Small fragments were immediately frozen in liquid nitrogen, and preserved at -80 °C. Total RNA extractions were performed with Maxwell 16 LEVsimplyRNA purification kits. RNA concentrations and quality were analyzed with an Agilent 2100 Bioanalyzer, and sequencing was done using paired-end in one lane of Illumina HiSeq 3000. RNA-seq libraries were prepared according to Illumina's protocols using the Illumina TruSeq Stranded mRNA sample prep kit to analyze mRNA. Briefly, mRNA was selected using poly-T beads. Then, RNA were fragmented to generate double stranded cDNA and adaptators were ligated to be sequenced. 11 cycles of PCR were applied to amplify libraries. Library quality was assessed using a Fragment

Published online: 05 July 2021

Analyser and libraries were quantified by qPCR using the Kapa Library Quantification Kit. RNA-seq experiments have been performed on an Illumina HiSeq3000 using a paired-end read length of 2 × 150 pb with the Illumina HiSeq3000 sequencing kits.

Marker development and screening

Sequence reads quality was assessed using FastQC v0.10.1 (Andrews 2010). Trimmomatic v0.32 was used to remove low quality reads with a Phred score below 20, as well as the Illumina adapters (Bolger et al. 2014). FastQC was performed again to verify the integrity of the remaining raw Illumina sequence reads. High-quality reads were then used for the de novo transcriptome assembly, using Trinity with its default *k-mer* value of 25 (Grabherr et al. 2011). DiscoSnp + + v2.2.10 (Uricaru et al. 2015) was then used with default parameters to call SNPs using the assembled transcriptome as a reference.

A total of 3,391 SNPs with good coverage (> 100 reads) were identified from the transcriptome of *L. sarmentosa*. A set of 320 SNPs were randomly selected and submitted for assay design using the MassARRAY Assay Designer version 4.0.0.2 (Agena Biosciences). The assay was performed on the genomic DNA extracted from 95 samples of *L. sarmentosa* collected in the NW Mediterranean (Table 1), using the DNeasy Blood and Tissue kits from QIAGEN. Allele calling was carried out with Typer Viewer v.4.0.24.71 (Agena Biosciences). Monomorphic SNPs, loci with weak or ambiguous signal (i.e., displaying more than three clusters of genotypes or unclear cluster delimitation) and loci with too much missing data were all discarded. The genetic parameters, including observed heterozygosity, expected heterozygosity, and the Hardy–Weinberg equilibrium (HWE), were calculated using GenoDive 2.0b23 (Meirmans and van Tienderen 2004).

Table 1 Sampling of *Leptogorgia sarmentosa* in the NW Mediterranean Sea

Site	Latitude	Longitude	Depth (m)	Date	n
Gruissan	43.1240	3.1645	10	07/2019	12
Valras	43.2329	3.3241	10	07/2019	15
Agde – point A	43.2520	3.5378	10	07/2019	15
Agde – point B	43.2543	3.4755	20	07/2019	15
Barcares	42.8240	3.0578	16	07/2019	15
Leucate	42.8954	3.0701	16	07/2019	10
Vendres	43.1934	3.2587	10	02/2020	13

Results

A total of 236 SNPs were successfully amplified across most of the samples. A subset of 153 loci were found to be polymorphic, resulting in a detectable rate of 64.83 %. The observed heterozygosity ranged from 0 to 0.54, while the expected heterozygosity varied from 0.00 to 0.49 (Supplementary Material). All 153 loci were found to be in HWE over all populations ($p > 0.01$). These results provide an invaluable resource for the future of population genetic studies of one of the most ubiquitous gorgonian species in the Mediterranean Sea.

Supplementary Information The online version of this article contains supplementary material available (https://doi.org/10.1007/s12686-021-01218-3).

Acknowledgements This work was funded by the Agence de l'Eau Rhône-Méditerranée-Corse under project ICONE - Impacts actuels et potentiels de la CONnectivité Ecologique ajoutée par les récifs artificiels sur la biodiversité fixée des substrats durs du Golfe du Lion (PI, K. Guizien, AAP 2016). The authors gratefully acknowledge the helpful assistance during sampling of S. Blouet and the staff of the Aire Marine Protégée Agatoise. We also thank the scuba divers, JC Roca and B Hesse, for their technical assistance for field sampling in Banyuls sur Mer. We are grateful to the Bio2Mar core facility (http://bio2mar.obs-banyuls.fr) for providing access to instrumentation for RNA extractions. RNAseq was performed at the GeT-PlaGe core facility, INRAE Toulouse. SNP genotyping was performed at the Genome Transcriptome Facility of Bordeaux (grants from ANR-10-EQPX-16) with the help of Laure Dubois.

Author contributions MP designed the study. MM, AD, and EG, performed the sequencing and genotyping. MP analyzed the data and wrote the manuscript.

Funding This work was funded by the Agence de l'Eau Rhône-Méditerranée-Corse under project ICONE - Impacts actuels et potentiels de la CONnectivité Ecologique ajoutée par les récifs artificiels sur la biodiversité fixée des substrats durs du Golfe du Lion (PI, K. Guizien, AAP 2016).

Data availability The data will be submitted to the European Variation Archive (http://www.ebi.ac.uk/eva/).

Declarations

Conflict of interest The authors have no conflicts of interest to declare.

References

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

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

Breedy O, Guzmán H (2007) A revision of the genus *Leptogorgia* Milne Edwards & Haime, 1857 (Coelenterata: Octocorallia: Gorgoniidae) in the eastern Pacific. Zootaxa 1419

Carpine C (1963) Contribution à la conaissance des Gorgones Holaxonia de la Méditerranée occidentale. Bull Inst Océanogr Monaco 1270:1–52

Carpine C, Grasshoff M (1975) Les Gorgonaires de la Méditerranée. Bull l'Institut Océanographique Monaco 71:1–140

Grabherr M, Haas B, Yassour M, Levin J, Thompson D, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol 2011 May 15(7):644–652. doi:https://doi.org/10.1038/nbt.1883

Kinzie R III (1973) The zonation of West Indian gorgonians. Bull Mar Sci 23:93–155

Meirmans P, van Tienderen P (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Molecular Ecology Resources 4(4):792–794

Poliseno A, Feregrino C, Sartoretto S, Aurelle D, Wörheide G, McFadden C, Vargas S (2017) Comparative mitogenomics, phylogeny and evolutionary history of *Leptogorgia* (Gorgoniidae). Mol Phylogenet Evol 115(2017):181–189

Starmans A, Gutt J, Arntz W (1999) Mega-epibenthic communities in Arctic and Antarctic shelf areas. Mar Biol 135:269–280

True M (1970) Étude quantitative de quattre peuplements sciaphiles sur substrat rocheux dans la région marseilles. Bull Inst Oceanogr Monaco 1410:1–48

Uricaru R, Rizk G, Lacroix V, Quillery E, Plantard O, Chikhi R, Lemaitre C, Peterlongo P (2014) Reference-free detection of isolated SNPs. Nucleic Acids Res. doi:https://doi.org/10.1093/nar/gku1187

Weinberg S (1978) Mediterranean octocorallian communities and the abiotic environment. Mar Biol 49:41–57

Weinberg S (1976) Revision of the common Octacorallia of the Mediterranean circalittoral. I Gorgonacea Beaufortia 24:63–104