Development of a set of SNP markers for population genetics of the sea rose, Pentapora fascialis

Mariana Padrón, M. Milhes, M. Massot, Erwan Guichoux, Katell Guizien

Development of a set of SNP markers for population genetics of the sea rose, *Pentapora fascialis*

Padrón, M[1]., Milhes M.[2], Massot, M.[3], Guichoux, E.[3] & Guizien K.[1]

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

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

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

Correspondence email: guizien@obs-banyuls.fr

Abstract

A set of single nucleotide polymorphisms (SNP) was developed from the transcriptome of 10 individuals of the sea rose *Pentapora fascialis* (Cheilostomata) collected in Banyuls-sur-Mer Bay (NW Mediterranean sea). 53042 putative SNPs were identified and mapped on a de novo assembled transcriptome. A selected set of 320 SNPs with highest coverage and uniquely mapped in the assembled transcriptome was tested using a MassARRAY System on 95 individuals sampled in natural rocks and artificial reefs distributed over a hundred kilometers in the NW Mediterranean (outisde Banyuls-sur-Mer Bay). A total of 177 SNPs were successfully genotyped and found to be polymorphic. Among these, 154 SNPs were in Hardy-Weinberg equilibrium over all samples, with significant linkage disequilibrium in only 21 pairs of SNPs. The newly developed loci will be a valuable tool for population genetics studies of this calcifying bryozoan species whose erect structure makes it an engineering species and a target for conservation plans.

Keywords

Pentapora fascialis, Single nucleotide polymorphism, MassARRAY, bryozoan

Introduction

The erect cheilostomata bryozoan *Pentapora fascialis* (Pallas, 1766) is a common sessile mega-benthic species with medium size ranging from 10 to 20 cm which distributes in the Mediterranean on hard rock substrates, cobble and boulder areas and other living species (gorgonians) down to 60 m or more (Zabala 1986, Lombardi et al. 2008). *P. fascialis* has a dome-like shape made of convoluted bright orange bilamina sheets. Considered as a synonym species of *P. foliacea* (North Atlantic) by some taxonomists (Hayward and Ryland 1999), *P. fascialis* (Mediterranean) was redescribed as a distinct species based on a phylogenetic analysis of skeletal morphological differences (Lombardi et al. 2010). The present day geographical distribution of the *P. foliacea–P. fascialis* species complex extends discontinuously from the Hebrides, English Channel, the Bay of Biscay, along the Portuguese coast to Morocco, and into the Mediterranean as far east as the Adriatic Sea (Gautier, 1962; Zabala, 1986; Hayward & Ryland, 1999).

This calcifying bio-constructor is an habitat-forming species of temperate and cold-water oceanic shelves which responds to the ocean climate variations (Wood et al. 2012, Lombardi et al. 2014, Fortunato et al. 2015, Pagès-Escolà et al. 2018). *P. fascialis* is a modular organism growing at a rate up to 3.5 cm y^{-1} in height through asexual budding process and estimates of growth rate suggest that *P. fascialis* colonies in the Mediterranean can live up to ten years old or more (Cocito et al. 1998a). Little is known about its reproduction, except the presence of ovicells, a reliable indicator of the sexual reproductive status. Colonies of *P. fascialis* as small as 2.8 cm have been recorded as having ovicells, suggesting sexual maturity occurs at two to three year old (Cocito *et al.*, 1998b). By analogy with the other cheilostomata bryozoan species, *Bugula neritina,* it is assumed to reproduce sexually every year with a pelagic lecitotrophic larva dispersing during less than one day (Keough &

Chernoff, 1987). Despite larval dispersal traits are actually unknown, species presence on artificial reefs deployed over soft-bottom substrates (Blouet et al. 2022) and high recruitment rates found on artificial substrates such as plastic nets (Pagès-Escolà et al. 2020) indicate efficient larval dispersal over a few kms at least. This process is essential in the resilience of a fragile species vulnerable to physical disturbances such as storms and scuba-diving activity (Cocito et al. 1998b, Sala et al. 1996).

The objective of the present study is to propose a set of Single Nucleotide Polymorphism markers that will be useful to confirm the morphological species differentiation between *P. fascialis* and *P. foliacea* and study *P. fascialis* population genetics and connectivity.

Material and Methods

Transcriptome sequencing

The transcriptome of *P. fascialis* was sequenced from the total RNA of ten individuals collected by SCUBA diving in sites a few kms apart between 10 and 20 meters depth in Banyuls sur Mer, France. While a short sampling distance from the laboratory ensured the best preservation conditions for RNA extraction, it may be a limit for finding Single Nucleotide Polymorphism useful for future connectivity studies. However, with an assumed limited dispersal potential, population genetics differentiation was expected at the scale of a few kms. Taking samples from the Mediterranean sea only ensured SNPs determination will apply specifically to *P. fascialis*. 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 Marker development and screening. 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 used with k-mer size of 61 and minimum coverage of 20 per read set to detect SNPs from the 10 different individual transcriptomes. Among the 53042 SNPs detected, a total of 21481 SNPs had a good coverage (> 100 reads) in all 10 individual transcriptomes and 44575 SNPs could be mapped on the transcriptome assembly, either in a unique or in multiple position. The 380 SNPs with higher coverage among the SNPs mapped at a unique position were selected and tested as candidate markers for the MassArray genotyping. Using Assay Designer version 4.0.0.2 (Agena Biosciences), with a mass range of 4000-9000Da, an amplicon length range of 70-130bp, an extension primer length range of 11-30 bases and a minimum analyte peak separation of 20 Da, four cost-effective multiplexes of 40 markers were developed for a total of 320 SNPs. The assay was performed on the genomic DNA extracted from 95 samples of *P. fascialis* 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).

Location	Site	Latitude (decimal degrees)	Longitude (decimal degrees)	Depth (m)	N
Cap d'Agde - natural rocks	A3	43.252095	3.524767	30	11
	A4	43.244578	3.519568	25	24
Cap d'Agde – artificial reef	AGD26	43.2386	3.4992	20	20
Barcarès – artificial reef	BAR51	42.8241	3.0578	16	10
Leucate – artificial reef	LEU47	42.8955	3.0701	16	10
Canet – artificial reef	CAN53	42.7216	3.0722	28	10
Saint Cyprien – artificial reef	CYP 54	42.6293	3.0720	30	10

Table 1.- Location and number of individuals (N) of *Pentapora fascialis* used for SNPs validation using MassArray

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. Genetics parameters, observed heterozygosity and expected heterozygosity were calculated using Genepop 4.7.5 (Rousset 2008). Hardy–Weinberg equilibrium (HWE) for each locus and Linkage Disequilibrium (LD) between each pair of loci (with Bonferroni correction) was tested using Genepop 4.7.5 (Rousset 2008).

Results

Across the 320 SNPs tested, 143 SNPs either did not amplify or were monomorphic, probably resulting from the limited spatial extent of the ten individuals used for searching putative SNPs (Supplementary Material Table S1). The average call rate of the 177 SNPs detected in more than half of the 95 samples and polymorphic was 97.4 %. The observed heterozygosity ranged from 0 to 1, while the expected heterozygosity varied from 0.0161 to

1, with 32 SNPs displaying an heterozygosity lower than 0.1 (highlighted in orange in the Supplementary Material Table S2). Among the 177 SNPs, 23 departed significantly from HWE proportions over the 95 samples spanning the 7 locations, (*p-value* < 0.05, highlighted in red in Supplementary Material Table S2). LD was rare, present in only 21 pairs of loci out of the 51040 pairs tested (the 32 loci involved are highlighted in yellow in Supplementary Material Table S2). These results provide a valuable resource for future population genetic studies of this endemic bryozoan species in the Mediterranean Sea.

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 (2017 7267 / 2017 7268 / 2018 0697, 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 and KG designed the study. MM, AD, and EG, performed the sequencing and genotyping. MP analyzed the data. KG wrote the manuscript. All authors reviewed 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 (2017 7267 / 2017 7268 / 2018 0697, PI K. Guizien, AAP 2016).

Data Availability The RNA seq data will be released publicly through National Center for Biotechnology Information portal in March 2024 (https://submit.ncbi.nlm.nih.gov).

References

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

Blouet S, Bramanti L, Guizien K (2022) Artificial reefs geographical location matters more than shape, age and depth for sessile invertebrate colonization in the Gulf of Lion (NorthWestern Mediterranean Sea). Peer Community Journal 2(article no e24). https://doi.org/10.24072/pcjournal.107.

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

Cocito S, Sgorbini S, Bianchi C (1998a) Aspects of the biology of the bryozoan *Pentapora fascialis* in the northwestern Mediterranean. *Marine Biology* **131**, 73–82 https://doi.org/10.1007/s002270050298

Cocito S., Ferdeghini F., & Sgorbini S. (1998b) *Pentapora fascialis* (Pallas) [Cheilostomata: Ascophora] colonization of one sublittoral rocky site after sea-storm in the northwest Mediterranean. *Hydrobiologia*, **375/376**, 59-66.

Eisenbarth S., Gehling M., Harder A., Steffan B. (2002) Pentaporins A, B and C: disulfides

169 from the marine bryozoan Pentapora fascialis, Tetrahedron, 58 (42): 8461-8464, ISSN 0040-
170 4020, https://doi.org/10.1016/S0040-4020(02)01050-5.

171 Fortunato H (2015) Bryozoans in climate and ocean acidification research: A reappraisal of
172 an under-used tool, Regional Studies in Marine Science 2: 32-44, ISSN 2352-4855,
173 https://doi.org/10.1016/j.rsma.2015.08.010.

174 Gautier YV (1962) Recherches écologiques sur les Bryozoaires cheilostomes en
175 Méditerranée occidentale. Thèse de l'Université d'Aix-Marseille.

176 Grabherr M, Haas B, Yassour M, Levin J, Thompson D, Amit I, Adiconis X, Fan L,
177 Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F,
178 Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length
179 transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol
180 2011 May 15(7):644–652. doi:https://doi.org/10.1038/nbt.1883

181 Hayward PJ, Ryland JS (1999) Synopses of the British Fauna (new series): Cheilostomatous
182 Bryozoa Part II: Hippothoidea-Celleporoidea. London: The Linnean Society of London and
183 Estuarine and Coastal Sciences Association.

184 Keough MJ, Chernoff H (1987) Dispersal and population variation in the bryozoan *Bugula*
185 *neritina*. *Ecology*, **68**, 199 - 210.

186 Lombardi C., Cocito S., Occhipinti-Ambrogi A., & Porter J. (2008). Distribution and
187 morphological variation of colonies of the bryozoan Pentapora fascialis (Bryozoa:
188 Cheilostomata) along the western coast of Italy. *Journal of the Marine Biological Association*
189 *of the United Kingdom*, *88*(4), 711-717. doi:10.1017/S0025315408001525

190 Lombardi C, Taylor PD, Cocito S (2010) Systematics of the Miocene–Recent bryozoan genus

191 Pentapora (Cheilostomata). Zoological Journal of the Linnean Society, 2010, 160, 17–39.

192 Lombardi C, Taylor PD, Cocito S (2014) Bryozoan constructions in a changing
193 Mediterranean Sea. In: Godofredo, S., Dubinsky, (Eds.), The Mediterranean Sea: Its History
194 and Present Challenges. Springer, Dordrecht, pp. 373–384.

195 Pagès-Escolà M, Hereu B, Garrabou J, Montero-Serra I, Gori A, Gómez-Gras D, Figuerola B,
196 Linares C (2018) Divergent responses to warming of two common co-occurring
197 Mediterranean bryozoans. Scientific Reports, 8, 17455. https://doi.org/10.1038/s41598-018-
198 36094-9

199 Pagès-Escolà M, Linares C, Gómez-Gras D, Medrano A, Hereu B. (2020) Assessing the
200 effectiveness of restoration actions for Bryozoans: The case of the Mediterranean Pentapora
201 fascialis. Aquatic Conserv: Mar Freshw Ecosyst. 30:8–19. https://doi.org/10.1002/aqc.3236

202 Rousset, F., 2008. Genepop'007: a complete reimplementation of the Genepop software for
203 Windows and Linux. Mol. Ecol. Resources 8: 103-106.

204 Sala E, Garrabou J, Zabala M (1996) Effects of diver frequentation on Mediterranean
205 sublittoral populations of the bryozoan *Pentapora fascialis* . *Marine Biology* **126**, 451–459.
206 https://doi.org/10.1007/BF00354627

207 Uricaru R, Rizk G, Lacroix V, Quillery E, Plantard O, Chikhi R, Lemaitre C, Peterlongo P
208 (2014) Reference-free detection of isolated SNPs. Nucleic Acids Res. doi:https://doi.org/10.
209 1093/nar/gku1187

210 Wood, A.C.L., Probert, P.K., Rowden, A.A., Smith, A.M., 2012. Complex habitat generated
211 by marine bryozoans: a review of its distribution, structure, diversity, threats and
212 conservation. Aquat. Conserv. 22, 547–563. http://dx.doi.org/10.1002/aqc.2236.

213 Zabala, M., 1986. In: Fauna dels briozous dels Països Catalans. Institut d'Estudis Catalans,

214 pp. 833.