



**HAL**  
open science

## Development of a set of SNP markers for population genetics of the sea peach (*Halocynthia papillosa*)

Mariana Padrón, Marine Milhes, Adline Delcamp, Erwan Guichoux, Katell Guizien

### ► To cite this version:

Mariana Padrón, Marine Milhes, Adline Delcamp, Erwan Guichoux, Katell Guizien. Development of a set of SNP markers for population genetics of the sea peach (*Halocynthia papillosa*). Conservation Genetics Resources, 2023, 10.1007/s12686-023-01313-7 . hal-04004739

**HAL Id: hal-04004739**

**<https://hal.science/hal-04004739v1>**

Submitted on 25 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

1 **Licence CC-BY**

2 **Development of a set of SNP markers for population genetics of the sea peach**

3 **(*Halocynthia papillosa*)**

4 Padrón, M.<sup>1</sup>, Milhes M.<sup>2</sup>, Delcamp, A.<sup>3</sup>, Guichoux, E.<sup>3</sup> & Guizien K.<sup>1</sup>

6 <sup>1</sup> Laboratoire d'Ecogéochimie des Environnements Benthique, CNRS, Sorbonne Université,

7 UMR 8222, F-66650, Banyuls-sur-Mer, France

8 <sup>2</sup> INRA, GeT-PlaGe, Genotoul, 31326 Castanet- Tolosan, France

9 <sup>3</sup> Univ. Bordeaux, INRAE, BIOGECO, F-33610 Cestas, France

10 Corresponding author: [guizien@obs-banyuls.fr](mailto:guizien@obs-banyuls.fr)

12 **Abstract**

13 A set of single nucleotide polymorphism (SNP) markers was developed from the  
14 transcriptome of the solitary ascidian *Halocynthia papillosa*. 1399 putative SNPs with  
15 coverage larger than 100 reads were identified, and a selected set of 320 SNPs was tested  
16 using a MassARRAY System on 95 samples from the NW Mediterranean. A total of 175  
17 SNPs were successfully genotyped, polymorphic and in Hardy-Weinberg equilibrium over all  
18 samples. The newly developed loci will be a valuable tool for population genetics studies of  
19 this endemic ascidian species of the Mediterranean Sea.

20 **Keywords**

21 *Halocynthia papillosa*, Single nucleotide polymorphism, MassARRAY

23 **Introduction**

24 The solitary ascidian *Halocynthia papillosa* is an endemic species of the Mediterranean sea  
25 (Pérès and Picard, 1964) and Eastern Atlantic from the Canary Islands to the Berlengas

26 Islands in Portugal (Ramos-Esplá, 1991). It is a common sessile mega-benthic species with  
27 medium size of about 15 cm that distributes evenly with low population density (a few  
28 individuals per meter square) on hard bottom between 10 to 100 m in the Mediterranean sea  
29 (Antoniadou et al. 2006; Coppari et al. 2014). Tunicates have been extensively studied for  
30 their tunique made of a unique animal cellulose (Helbert et al. 1998). In contrast, the ecology  
31 of the species has driven less interest. The species is a filter-feeder with a broad natural diet  
32 including detrital organic matter, heterotrophic bacteria, *Prochlorococcus* sp., *Synechococcus*  
33 sp., protozoa and phytoplankton with a mean size range from 0.6 to 70  $\mu\text{m}$  (Ribes et al.  
34 1998). It is a simultaneous hermaphrodite with a spawning period extending over 2 to 3  
35 months, between September and November (Becerro and Turon, 1992). Ecologically, it is  
36 considered an indicator of good conditions and unperturbed rocky areas (Naranjo et al. 1996).  
37 Low population density of *H. papillosa* has been correlated to high SCUBA diving frequency  
38 (Luna-Pérez et al., 2010). The deleterious impact of SCUBA diving on *H. papillosa* was  
39 further tested experimentally by applying treatments with various intensities of flapping, with  
40 highest flapping intensity altering *H. papillosa* growth and abundance (Luna-Pérez et al.,  
41 2011). Nevertheless, the species has been found on artificial reefs deployed over soft-bottom  
42 substrates, showing the potential of the species to survive in turbid environments (Blouet et  
43 al. 2022). In addition, the colonisation of artificial reefs located at a few kms from any rocky  
44 substrate question the dispersal potential of this species, an essential mechanism for  
45 population recovery in a sensitive species. The objective of the present study is to propose a  
46 set of Single Nucleotide Polymorphism markers that will be useful to study *H. papillosa*  
47 population genetics and connectivity.

48

## 49 **Material and Methods**

### 50 *Transcriptome sequencing*

51 The transcriptome of *H. papillosa* was sequenced from the total RNA of ten individuals  
52 collected by SCUBA diving between 10 and 20 meters depth in Banyuls sur Mer, France.  
53 Small fragments were immediately frozen in liquid nitrogen, and preserved at -80 °C. Total  
54 RNA extractions were performed with Maxwell 16 LEV simplyRNA purification kits. RNA  
55 concentrations and quality were analysed with an Agilent 2100 Bioanalyzer. RNA-seq  
56 libraries were prepared according to Illumina's protocols using the Illumina TruSeq Stranded  
57 mRNA sample prep kit, using 11 cycles of PCR to amplify libraries. Library quality was  
58 assessed using a Fragment Analyzer. Libraries were quantified by qPCR using the Kapa  
59 Library Quantification Kit. Sequencing was achieved on an Illumina HiSeq3000 using a  
60 paired-end read length of  $2 \times 150$  pb.

61

#### 62 *Marker development and screening*

63 Sequence read quality was assessed using FastQC v0.10.1 (Andrews 2010). Trimmomatic  
64 v0.32 was used to remove low quality reads with a Phred score below 20, as well as the  
65 Illumina adapters (Bolger et al. 2014). FastQC was performed again to verify the integrity of  
66 the remaining raw Illumina sequence reads. High-quality reads were then used for the de  
67 novo transcriptome assembly, using Trinity with its default k-mer value of 25 (Grabherr et al.  
68 2011). DiscoSnp ++ v2.2.10 (Uricaru et al. 2015) was used independently with kmer size of  
69 61 and a minimal coverage per read 20 to call SNPs using the high-quality reads of the 10  
70 individuals transcriptomes. A total of 28320 SNPs were found among which 1399 SNPs had  
71 good coverage ( $> 100$  reads) in each of the 10 individual's transcriptomes. A set of 320 SNPs  
72 were randomly selected and submitted for assay design using the MassARRAY Assay  
73 Designer version 4.0.0.2 (Agena Biosciences). The assay was performed on the genomic  
74 DNA extracted from 95 samples of *H. papillosa* collected in 6 locations in the NW  
75 Mediterranean (Table 1, 5 locations in Agde and 1 location in Saint-Cyprien), using the

76 DNeasy Blood and Tissue kits from QIAGEN. Allele calling was carried out with Typer  
77 Viewer v.4.0.24.71 (Agena Biosciences). Monomorphic SNPs, loci with weak or ambiguous  
78 signal (i.e., displaying more than three clusters of genotypes or unclear cluster delimitation)  
79 and loci with too much missing data were all discarded. The genetic parameters, including  
80 observed heterozygosity, expected heterozygosity, and the Hardy–Weinberg equilibrium  
81 (HWE), were calculated using Genepop 4.7.5 (Rousset 2008).

82 **Table 1.-** Location and number of individuals (N) of *Halocynthia papillosa* used for SNP  
 83 validation using MassArray  
 84

Location	Site	Latitude (decimal degrees)	Longitude (decimal degrees)	Depth (m)	N
Cap d'Agde - natural rocks	A1	43,252556	3,519840	13	20
	A4	43,244578	3,519568	25	15
Cap d'Agde – artificial reef	AGD26	43,2386	3,4992	20	10
	AGD28	43,2521	3,5378	20	6
	AGD30	43,2543	3,4756	10	20
Saint Cyprien – artificial reef	CYP 54	42,6293	3,0720	30	20

85

86

## 87 **Results**

88 A total of 220 SNPs were successfully amplified across more than half of the samples. A  
 89 subset of 188 loci were found to be polymorphic, resulting in a detectable rate of 85.4 %. The  
 90 observed heterozygosity ranged from 0.011 to 0.989, while the expected heterozygosity  
 91 varied from 0.011 to 0.503 (Supplementary Material). Among the 188 loci, 175 were found in  
 92 HWE proportions over 95 samples ( $p > 0.01$ ). These results provide a valuable resource for  
 93 future population genetic studies of this endemic ascidian species in the Mediterranean Sea.

94

## 95 **Acknowledgements**

96 This work was funded by the Agence de l'Eau Rhône-Méditerranée-Corse under project  
 97 ICONE - Impacts actuels et potentiels de la CONnectivité Ecologique ajoutée par les récifs  
 98 artificiels sur la biodiversité fixée des substrats durs du Golfe du Lion (PI, K. Guizien, AAP  
 99 2016). The authors gratefully acknowledge the helpful assistance during sampling of S.  
 100 Blouet and the staff of the Aire Marine Protégée Agatoise. We also thank the scuba divers, JC  
 101 Roca and B Hesse, for their technical assistance for field sampling in Banyuls sur Mer. We  
 102 are grateful to the Bio2Mar core facility (<http://bio2mar.obs-banyuls.fr>) for providing access  
 103 to instrumentation for RNA extractions. RNAseq was performed at the GeT-PlaGe core

104 facility, INRAE Toulouse. SNP genotyping was performed at the PGTB  
105 (doi:10.15454/1.5572396583599417E12) with the help of Laure Dubois.

106

107 Author contributions: MP and KG designed the study. MM, AD, and EG, performed the  
108 sequencing and genotyping. MP analyzed the data. KG wrote the main manuscript text. All  
109 authors reviewed the manuscript.

110

111 Funding: This work was funded by the Agence de l'Eau Rhône-Méditerranée-Corse under  
112 project ICONE - Impacts actuels et potentiels de la CONnectivité Ecologique ajoutée par les  
113 récifs artificiels sur la biodiversité fixée des substrats durs du Golfe du Lion (PI, K. Guizien,  
114 AAP 2016, grant numbers : 2017 7267 / 2017 7268 / 2018 0697).

115

116 Data availability: The RNA seq data will be released publicly through National Center for  
117 Biotechnology Information portal (<https://submit.ncbi.nlm.nih.gov>).

118

## 119 **References**

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

122 Antoniadou C, Voultziadou E, Chariton Chintiroglou C (2006) Sublittoral megabenthos along  
123 cliffs of different profile (Aegean Sea, Eastern Mediterranean). Belg. J. Zool., 136 (1) : 69-79

124 Becerro MA, Turon X (1992) Reproductive cycles of the ascidians *Microcosmus sabatieri*  
125 and *Halocynthia papillosa* in the Northwestern Mediterranean. PSZN I: Mar Ecol 13:363-373

126 Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina  
127 Sequence Data. Bioinformatics 30(15):2114–2120. <https://doi.org/10.1093/bioinforma>

128 tics/btu170

129 Blouet S, Bramanti L, Guizien K. (2022) Artificial reefs geographical location matters more  
130 than shape, age and depth for sessile invertebrate colonization in the Gulf of Lion  
131 (NorthWestern Mediterranean Sea). Peer Community Journal 2 (article no. e24) doi :  
132 10.24072/pcjournal.107. <https://peercommunityjournal.org/articles/10.24072/pcjournal.107/>

133 Coppari, M., Gori, A., Rossi, S. (2014) Size, spatial, and bathymetrical distribution of the  
134 ascidian *Halocynthia papillosa* in Mediterranean coastal bottoms: benthic–pelagic coupling  
135 implications. *Mar Biol* **161**, 2079–2095. <https://doi.org/10.1007/s00227-014-2488-5>

136 Grabherr M, Haas B, Yassour M, Levin J, Thompson D, Amit I, Adiconis X, Fan L,  
137 Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F,  
138 Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length  
139 transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol*  
140 2011 May 15(7):644–652. doi:<https://doi.org/10.1038/nbt.1883>

141 Helbert W, Nishiyama Y, Okano T, Sugiyama J (1998) Molecular imaging of *Halocynthia*  
142 *papillosa* cellulose. *Journal of Structural Biology* 124 (1): 42-50,  
143 <https://doi.org/10.1006/jsbi.1998.4045>.

144 Luna-Pérez B., Valle-Pérez C., Vega Fernandez T., Sanchez-Lizaso J.L., Ramos-Esplà A.

145 A. (2010) *Halocynthia papillosa* (Linnaeus, 1767) as an indicator of SCUBA diving impact.  
146 *Ecol. Ind.* 10, 1017–1024

147 Luna-Pérez B, Valle-Pérez C, Sánchez-Lizaso J.L. (2010) *Halocynthia papillosa* as SCUBA  
148 diving impact indicator: An in situ experiment, *Journal of Experimental Marine Biology and*  
149 *Ecology* 398 (1–2): 33-39, <https://doi.org/10.1016/j.jembe.2010.11.013>.

150 Naranjo SA, Carballo JL, García-Gómez JC (1996)



151 Effects of environmental stress on ascidian populations in Algeciras Bay (southern Spain).  
152 Possible marine bioindicators? Mar Ecol Prog Ser 144:119-131 doi:10.3354/meps144119

153 Pérès, J.M., Picard, J., 1964. Nouveau manuel de bionomie benthique de la mer  
154 Méditerranée. Rec. Trav. Sta. Mar. Endoume. 31, 5–137.

155 Ramos-Esplá, A.A., 1991. Ascidijs litorales del Mediterráneo ibérico. Faunística, ecología  
156 y biogeografía. Universidad de Alicante Press, Alicante.

157 Ribes, M., Coma, R., Gili, J.M., 1998. Seasonal variation of in situ rates by the temperate  
158 ascidians *Halocynthia papillosa*. Mar. Ecol. Progr. Ser. 175, 201–213.

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

161 Uricaru R, Rizk G, Lacroix V, Quillery E, Plantard O, Chikhi R, Lemaitre C, Peterlongo P  
162 (2014) Reference-free detection of isolated SNPs. Nucleic Acids Res. doi:[https://doi.org/10.](https://doi.org/10.1093/nar/gku1187)  
163 [1093/nar/gku1187](https://doi.org/10.1093/nar/gku1187)

164

165

166