Isolation and characterization of nine microsatellite loci of Terapon jarbua (Forsskål, 1775) from Socotra Island (Gulf of Aden) using multiplex PCR

Edouard Lavergne, Isabelle Calves, Uwe Zajonz, Jean Laroche

Title:

Isolation and characterization of nine microsatellite loci of *Terapon jarbua* (Forsskål, 1775) from Socotra Island (Gulf of Aden) using multiplex PCR.

Author's names with initials:

E. LAVERGNE [1, 2, 3] (EL)

I. CALVES [3] (IC)

U. ZAJONZ [2, 1] (UZ)

J. LAROCHE [3] (JL)

Full postal addresses:

[1] Senckenberg Forschungsinstitut und Naturmuseum, Sektion Ichthyologie – Senckenberganlage 25, 60325 Frankfurt am Main, Deutschland

[2] Biodiversität und Klima Forschungszentrum (BiK-F) – Senckenberganlage 25, 60325 Frankfurt am Main, Deutschland

[3] UMR 6539 CNRS/UBO/IRD - LEMAR (Laboratoire des Sciences de l'Environnement Marin), IUEM (Institut Universitaire Européen de la Mer), Université de Bretagne Occidentale – Technopôle Brest Iroise, Rue Dumont d'Urville 29280 Plouzané, France

Keywords: *Terapon jarbua*, microsatellites, population genetics, Socotra Island

Corresponding author:

Edouard Lavergne

Senckenberg Forschungsinstitut und Naturmuseum, Sektion Ichthyologie – Senckenberganlage 25, 60325 Frankfurt am Main, Deutschland

Fax: +49 69 75 42 12 53

E-mail: edouard.lavergne@senckenberg.de

Abstract

Ten polymorphic microsatellite loci were identified and characterized from 22 individuals of *Terapon jarbua* (Forsskål, 1775) from Socotra Island (Gulf of Aden, Yemen). Microsatellite polymorphism was tested, revealing 4 to 19 alleles per locus. The observed heterozygosity values ranged from 0.318 to 0.909. Nine loci out of ten conformed to Hardy-Weinberg proportions. They did not show evidence for null alleles and gametic disequilibrium. These loci will be used in an ongoing study of the population structure of this species; associated with a study assessing habitat connectivity based on otolith microchemistry of *T. jarbua*. Results are expected to inform estuarine conservation efforts on Socotra Island and in the Gulf of Aden region.

Terapon jarbua (Terapontidae, Perciformes) inhabits marine and brackish waters of the Indo-Pacific, from the Red Sea and East coast of Africa to Samoa. Its juveniles have been observed to thrive even in coastal freshwater courses. The type locality of the species is Jeddah, Saudi Arabia, and the type series is deposited in the Zoological Museum of Copenhagen (Klausewitz & Nielsen 1965, Nielsen 1974). Although advances have been made in the taxonomy and biology of the Terapontidae (Vari 1978, Whitfield & Blaber 1978, Miu 1990),

no in-depth studies of the reproductive ecology, life history strategy and population structure of *T. jarbua* have been conducted as of yet. The microsatellite markers presented herein are therefore the first developed for this species, and will be especially supportive to further genetic studies of the species.

DNA libraries for *Terapon jarbua* enriched for microsatellite sequences containing AAC, ATG, CATC and TAGA repeat motifs were constructed by Genetic Identification Services following the method described by Jones et al. (2002). Resulting recombinant clones were selected at random and sequenced on an Applied Biosystems™ 377 DNA Sequencer, using Amersham's DYEnamic™ ET Terminator Cycle Sequencing Kit (Amersham Biosciences P/N US81050). Initial polymerase chain reaction (PCR) primers were designed for flanking regions of microsatellite containing sequences using DESIGNER PCR v1.03 (Research Genetics Inc.).

Genomic DNA was extracted from muscle tissue, preserved in 95% ethanol of 22 individuals collected in March 2007 by seine net at Matief Estuary (12° 26' 48.5'' N and 54° 18' 17.6'' E) on Socotra Island (Fig. 1). Extractions were performed using AcroPrep™ 96 well Filter Plates (1mL) with 1µm Glass Fiber media (PALL® 5051), following the extraction protocol for DNA barcoding by Ivanova et al. (2006). PCRs conducted on a GeneAmp PCR system 9700 (Applied Biosystems™) were optimized for each primer individually on four randomly selected samples. After optimisation, PCR reactions were multiplexed in a total reaction volume of 10µL, using 5µL of Master-mix, 1µL of Solution Q both from the Qiagen Multiplex PCR kit, 1µL of multiplexed primer-mix (Table 1) containing specific primers and the labelled universal primers 6Fam-TAGTCGACGACCGTTA, Yakima Yellow YY-TCGGATAGCTAGTCGT, and Dargonfly Orange DO-CTGGCCGTCGTTTTAC (Chang et al. 2004) in order to avoid the expenses of using specific fluorescent primers, 1µL of template

DNA (30-50 ng.µL^{-1}) and 2µL of water. Touchdown PCR conditions consisted of an initial denaturing step at 95°C (15') followed by 11 cycles at 94°C (30''), 63-53°C (1'30'') and 72°C (1'), followed by 25 cycles at 94°C (30''), 53°C (1'30'') and 72°C (1'). A final elongation step at 60°C (30 min) ended the PCR. 3µL of PCR products were added to 12µL of formamide and 0.2µL of a 50-500bp size standard (GeneScan™-500 LIZ™) to visualise microsatellite alleles using an ABI 3130 Genetic Analyzer (Applied Biosystems™). Alleles were then scored using GeneMapper® Software v4.0 (Applied Biosystems™). The number of alleles and the observed and expected heterozygosity values were calculated using GENETIX v4 (Belkhir 2004); deviation from the Hardy–Weinberg proportions (Fisher's exact test) and gametic disequilibrium (Fisher's exact test) among loci were tested using GENEPOP v4 (Raymond & Rousset 1995, Rousset 2008). Both tests were corrected for multiple simultaneous tests by calculating the q-value of each test which measures the minimum *false discovery rate (FDR)* that is incurred when calling that test significant. The bootstrap method was chosen as recommended by the authors for a limited number of p-values (Storey 2002). The q-values were calculated using the R package QVALUE (www.r-project.org, Ihaka & Gentleman 1996, Storey 2002, Storey & Tibshirani 2003, Storey 2003, Storey et al. 2004). This correction was preferred over the commonly used sequential Bonferroni correction (Rice 1989) following Moran (2003). Null allele frequencies were calculated based on Brookfield (1996) using the program MICRO-CHECKER (Van Oosterhout et al. 2004).

Ten out of 12 loci were reliably amplified and found to be polymorphic for *T. jarbua* (Table 2). The number of alleles per locus ranged from 4 to 19, with observed and expected heterozygosity values varying respectively from 0.318 to 0.909, and from 0.328 to 0.941. Exact tests after correction indicated that one locus (B107) deviated significantly from Hardy-Weinberg proportions (q-value < 0.01). Exact test for gametic disequilibrium yielded five

97 weakly significant p-values (p-value < 0.05) out of 45 pair wise comparisons; none of which
98 were ultimately found to be significant following the FDR correction (q-value > 0.05). No
99 locus showed evidence for a null allele. Therefore, nine of the ten markers presented in this
100 study can be applied in studying the genetic structure of *T. jarbua* populations. Such studies
101 are expected to be instrumental in future estuary conservation and management efforts (1) in
102 Yemen including the Socotra Archipelago, representing a UNESCO World Heritage (2008),
103 which is a treasure of marine biodiversity of regional and global importance (Zajonz & Krupp
104 2006), and (2) in the wider Indian Ocean.
105
106 **Acknowledgments**
107 The present study was mainly funded by the Total Foundation, Paris. It was partially
108 conducted at the Socotra Field Research Station and the Molecular Laboratory of the
109 Biodiversity and Climate Research Centre (BiK-F) based on financial support of the research
110 funding programme “LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-
111 ökonomischer Exzellenz" of Hesse's Ministry of Higher Education, Research, and the Arts.
112 Additional mobility funds were provided by scholarships from the GRADE - Goethe Graduate
113 Academy of the Goethe University Frankfurt a.M. and the International Doctoral College of
114 the European University of Brittany (UEB). A European Community fund (CPER-FEDER
115 project) allowed the LEMAR laboratory to acquire the high performance DNA sequencer used
116 in this study. We would like to express our gratitude to the Ministry of Water and
117 Environment and the Environment Protection Authority of Yemen for granting research and
118 sample export permits. The Socotra Conservation and Development Program (SCDP-UNDP)
119 is thanked for providing indispensable support during the field work. We wish to express our
120 gratitude to Fouad Naseeb and Motea Sheikh Aideed for their efficient assistance in the field,

to Dr. Le Maréchal for providing us the universal primer sequences and to Dr. Leila Meistertzheim for her useful comments on the manuscript. Dr. Friedhelm Krupp is cordially acknowledged for his support to the PhD project of the first author.

References

Belkhir K, Borsa P, Chikhi L, Raufaste N & Bonhomme F (1996-2004) GENETIX v4.05. logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Populations. Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, **5**, 453-455.

Chang EH, Menezes M, Meyer NC, Cucci RA, Vervoort VS, Schwartz CE & Smith RJ (2004) Branchio-oto-renal syndrome: The mutation spectrum in EYA1 and its phenotypic consequences. *Human Mutation*, **23**, 582-589.

Ihaka R & Gentleman R (1996) R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299-314.

Ivanova NV, de Waard J, Hebert PDN (2006) An inexpensive. automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes*, **6**, 998-1002.

Jones K, Levine K & Banks J (2002) Characterization of 11 polymorphic tetranucleotide microsatellites for forensic applications in California elk (*Cervus elaphus canadensis*). *Molecular Ecology Notes*, **2**, 425-427.

Klausewitz W & Nielsen J (1965) On Forsskål's collection in the Zoological Museum of Copenhagen. Spolia zoolgical museum Hauniensis, Copenhagen, pp.1-29.

Miu TS, Lee SC & Tzeng WN (1990) Reproductive Biology of *Terapon jarbua* from the Estuary of Tamshui River. *Journal of the Fishery Society of Taiwan*, **17(1)**, 9-20.

145 Moran MD (2003) Arguments for rejecting the sequential Bonferroni in ecological studies.
146 *Oikos*, **100(2)**, 403-405.
147 Nielsen JG (1974) Fish types in the Zoological Museum of Copenhagen. Zoological Museum,
148 University of Copenhagen, Denmark, 115 p.
149 Raymond M & Rousset F (1995) GENEPOP Version 1.2: population genetics software for exact
150 tests and ecumenicism. *Journal of Heredity*, **86**, 248-249.
151 Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223-225.
152 Rousset F (2008) GENEPOP'007: a complete reimplementation of the GENEPOP software for
153 Windows and Linux. *Molecular Ecology Resources*, **8**, 103-106.
154 Storey JD (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical*
155 *Society, Series B*, **64**, 479-498.
156 Storey JD & Tibshirani R (2003) Statistical significance for genome-wide experiments.
157 *Proceeding of the National Academy of Sciences*, **100**, 9440-9445.
158 Storey JD (2003) The positive false discovery rate: A Bayesian interpretation and the qvalue.
159 *Annals of Statistics*, **31**, 2013-2035.
160 Storey JD, Taylor JE & Siegmund D (2004) Strong control, conservative point estimation and
161 simultaneous conservative consistency of false discovery rates: A unified approach. *Journal*
162 *of the Royal Statistical Society, Series B*, **66**, 187-205.
163 Van Oosterhout C, Hutchinson WF, Wills DPM & Shipley P (2004) MICRO-CHECKER:
164 software for identifying and correcting genotyping errors in microsatellite data. *Molecular*
165 *Ecology Notes*, **4**, 535-538.
166 Vari RP (1978) The terapon perches (Percoidei. Teraponidae). A cladistic analysis and
167 taxonomic revision. Bulletin of the American Museum of Natural History, 159 (5), 175-340.

168 Whitfield AK & Blaber SJM (1978) Scale-eating habits of the marine teleost *Terapon jarbua*
169 (Forsskål. 1775). *Journal of Fish Biology*, **12**, 61-70.
170 Zajonz U & Krupp F (2006) Fish biogeography of Socotra. In: Socotra – A natural history of
171 the islands and their people. Cheung C & De Vantier L (Eds.), Van Damme K (Science Ed.),
172 Odyssey Books and Guides, Airphoto International Ltd, 408 pp.

173 **Figure 1: Map of Socotra Archipelago and sampling site location (Matief)**

Table1: Multiplex Primer-mix

Multiplex PCR 1	Multiplex PCR 2
5µL R C102 primer, 100nM	5µL R A4 primer, 100nM
5µL R C108 primer, 100nM	5µL R B107 primer, 100nM
5µL R C3 primer, 100nM	5µL R D108 primer, 100nM
5µL F C102 primer + univ. ext.,10nM	5µL F A4 primer + univ. ext., 10nM
5µL F C108 primer + univ. ext., 10nM	5µL F B107 primer + univ. ext., 10nM
5µL F C3 primer + univ. ext., 10nM	5µL F D108 primer + univ. ext., 10nM
5µL 6FAM-univ. primer, 100nM	5µL 6FAM-univ. primer, 100nM
5µL YY-univ. primer, 100nM	5µL YY-univ. primer, 100nM
5µL DO-univ. primer, 100nM	5µL DO-univ. primer, 100nM
117µL H_2O	117µL H_2O
150 µL	150µL

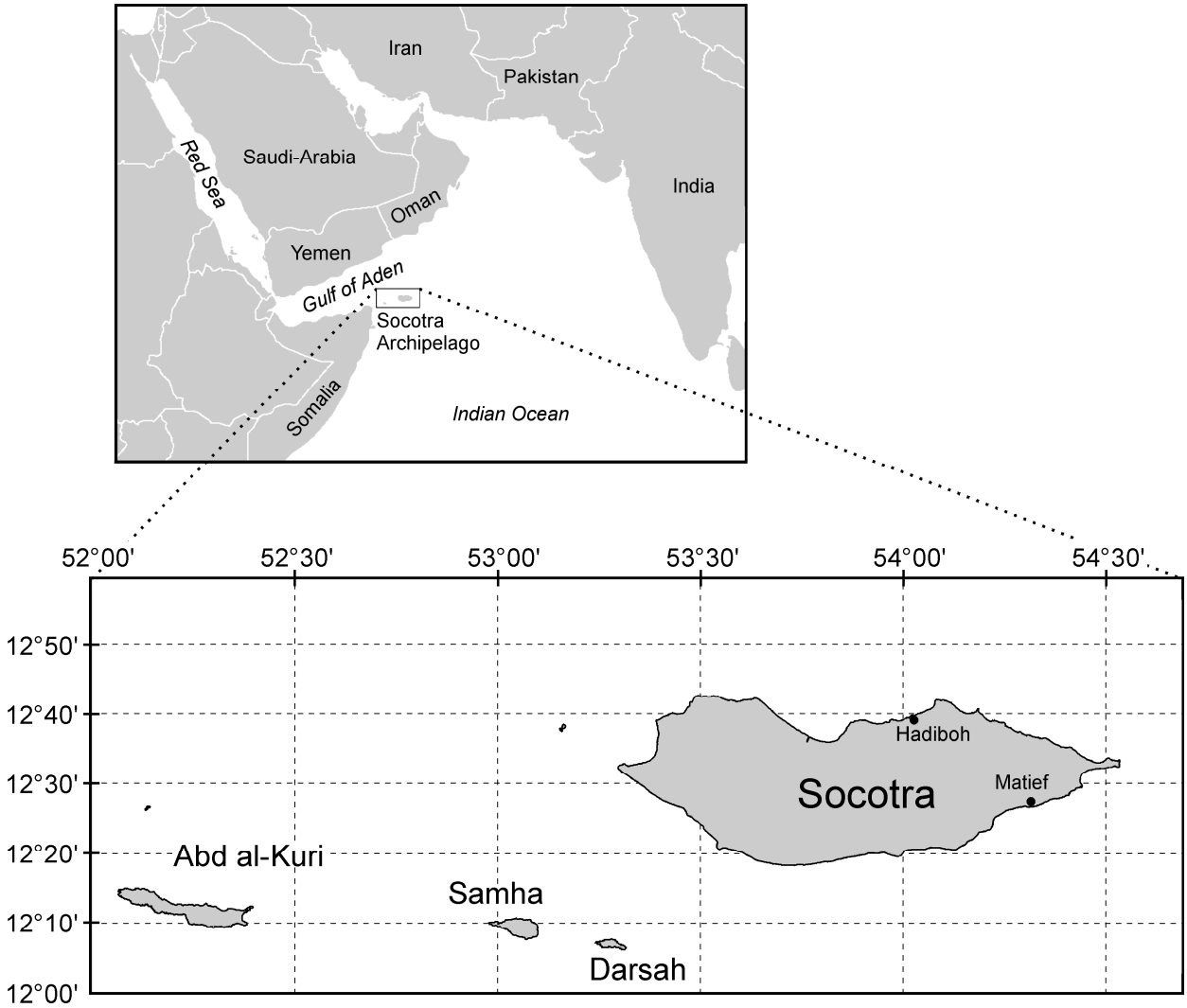
Multiplex PCR 3	Multiplex PCR 4
5µL R B103 primer, 100nM	5µL R B106 primer, 100nM
5uL R C103 primer, 100nM	5µL R C105 primer, 100nM
5µL R D102 primer, 100nM	5µL R D3 primer, 100nM
5µL F B103 primer + univ. ext., 10nM	5µL F B106 primer + univ. ext., 10nM
5µL F C103 primer + univ. ext., 10nM	5µL F C105 primer + univ. ext., 10nM
5µL F D102 primer + univ. ext., 10nM	5µL F D3 primer + univ. ext., 10nM
5µL 6FAM-univ. primer, 100nM	5µL 6FAM-univ. primer, 100nM
5µL YY-univ. primer, 100nM	-
5µL DO-univ. primer, 100nM	7,5µL DO-univ. primer, 100nM
117µL H_2O	119,5µL H_2O
150µL (Total volume)	150µL (Total volume)

Primer and universal extension association is shown in Table 2

Table 2: Primer sequences for 10 microsatellite loci and allele statistics in one population (N = 22) of *Terapon jarbua*

Loci	EMBL #	Repeat motif	5' Universal Extension Primers 5' → 3'	*i*	A	R	H_E	H_O	P_{HW}	Q_{HW}
C102	FR719958	$(CTAT)_{13}(CCAT)_{28}(CTAT)_2$	F: CTGGCCGTCGTTTTACGTCTCCCTCCCTCATGTCTG	1	19	171-257	0.938	0.909	0.0662	0.0513
			R: TTGCCACAGTGGACCTGTAG							
C108	FR719959	$(ATCC)_9ATTT(ATCC)_2$	F: TCGGATAGCTAGTCGTCCATCCATTCATCCATCTAC	1	7	274-336	0.763	0.727	0.2318	0.0773
			R: GCTTTGGAGTATTTTGCAGTT							
C3	FR719960	$(CATC)_7$	F: TAGTCGACGACCGTTACATAATGAGCGAGGTCAGAT	1	6	278-305	0.791	0.727	0.0770	0.0513
			R: ATCACGGAGGTTCTAAGAGTC							
A4	FR719961	$(AAC)_{14}$	F: TCGGATAGCTAGTCGTACCTGCCTACTACAGCCTCAG	2	5	262-271	0.711	0.636	0.2146	0.0773
			R: CACTCCACTTGCCCATTTC							
B107	FR719962	$(CAT)CAA(CAT)_7$	F: CTGGCCGTCGTTTTACCCAAGTTCCTGATGCTAAAAG	2	10	205-231	0.868	0.773	0.003*	0.0008*
			R: AGACGATGATGGGATTATTTG							
B103	FR719963	$(CAT)_2CTT(CAT)_9CAC(CAT)$	F: TCGGATAGCTAGTCGTGGGCTGTAACAGTATGCAATG	3	4	215-227	0.669	0.727	0.5839	0.1557
			R: ATGCAGCACCTTCAGAGTTTA							
C103	FR719964	$(TCCA)_8$	F: CTGGCCGTCGTTTTACCTTTCAATAGCCAGGACTACC	3	4	183-198	0.665	0.545	0.1209	0.0645
			R: TCTTCCACACTGAGACTGCT							
B106	FR719965	$(CAT)_2CA(CAT)_4CAC(CAT)$	F: CTGGCCGTCGTTTTACAGAGGAGGACCACATAAACAC	4	11	112-159	0.859	0.727	0.0540	0.0513
			R: TTCCACCAGATGAGAGGAG							
C105	FR719966	$(ATCC)_5CTCC(ATCC)_{12}$	F: CTGGCCGTCGTTTTACAGCTTTGTGAGGCTAATACCAG	4	6	242-314	0.328	0.318	0.5475	0.1557
			R: AAGTCTTCTTCAACCCTGTGAG							
D3	FR719967	$(TCTA)_{10}$	F: TAGTCGACGACCGTTACAGTCCAGTAATGTCGTTTGT	4	19	283-349	0.941	0.909	0.2256	0.0773
			R: AGTGTTAGACAGGAGCACATG							

i, multiplex PCR index (Table 1); A, allele nb; R, size range; H_E, expected heterozygosity; H_O, observed heterozygosity; P_{HW}, p-value of HW exact test; Q_{HW}, q-value of HW exact test; *, Significant deviation from HWE.


Iran
Pakistan
Saudi-Arabia
Red Sea
India
Oman
Yemen
Gulf of Aden
Socotra
Archipelago
Somalia
Indian Ocean
52°00'
52°30'
53°00'
53°30'
54°00'
54°30'
12°50'
12°40'
12°30'
12°20'
12°10'
12°00'
Hadiboh
Socotra
Matief
Abd al-Kuri
Samha
Darsah

174