ICES Journal of Marine Science

ICES International Council for the Exploration of the Sea
CIEM Conseil International pour l'Exploration de la Mer

ICES Journal of Marine Science (2018), 75(4), 1465–1472. doi:10.1093/icesjms/fsx245

Original Article

Genetic analysis reveals harvested *Lethrinus nebulosus* in the Southwest Indian Ocean comprise two cryptic species

Amy J. E. Healey[1], Gavin Gouws[2], Sean T. Fennessy[3], Baraka Kuguru[4], Warwick H. H. Sauer[5], Paul W. Shaw[1], and Niall J. McKeown[1]*

[1]*Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Ceredigion, Wales SY23 3DA, UK*
[2]*South African Institute for Aquatic Biodiversity, Grahamstown 6139, South Africa*
[3]*Oceanographic Research Institute, South African Association for Marine Biological Research, Durban 4056, South Africa*
[4]*Tanzania Fisheries Research Institute (TAFIRI), PO Box 9750, Dar Es Salaam, Tanzania*
[5]*Department of Ichthyology & Fisheries Science, Rhodes University, Grahamstown, South Africa*

**Corresponding author: tel: +44(0)1970621986; fax: +44(0)1970622350; e-mail: njm2@aber.ac.uk*

Healey, A. J. E., Gouws, G., Fennessy, S. T., Kuguru, B., Sauer, W. H. H., Shaw, P. W., and McKeown, N. J. Genetic analysis reveals harvested *Lethrinus nebulosus* in the Southwest Indian Ocean comprise two cryptic species. – ICES Journal of Marine Science, 75: 1465–1472.

Received 12 October 2017; revised 9 December 2017; accepted 11 December 2017; advance access publication 5 January 2018.

This study initially aimed to investigate the genetic population/stock structuring of *Lethrinus nebulosus* in the Southwest Indian Ocean (SWIO) to inform management practices in light of emerging evidence of overharvesting of this species throughout its distribution. Adult samples were genotyped for 14 nuclear microsatellites and by sequencing fragments of the mtDNA control region and COI gene. A salient feature of the data was the congruent cyto-nuclear partitioning of samples into two high divergent, reciprocally monophyletic groups. This indicates that despite no *a priori* evidence, hitherto described *L. nebulosus* in the SWIO comprises two cryptic species that co-occur among southern samples. This intermingling indicates that, at least in southern samples, both species are being indiscriminately harvested, which may severely compromise sustainability. Limited microsatellite differentiation was detected within both species, though there was some evidence of isolation in the Mauritian population. In contrast, mtDNA revealed a pattern consistent with chaotic genetic patchiness, likely promoted by stochastic recruitment, which may necessitate a spatial bet-hedging approach to management to satisfy fishery management and conservation goals.

Keywords: biodiversity, cryptic species, microsatellite, mitochondrial DNA, stock, sustainability

Introduction

Genetic studies have yielded many insights into marine biodiversity, with important findings including the detection of significant genetic population structuring (Shaw *et al.*, 1999; Knutsen *et al.*, 2011; McKeown *et al.*, 2015) and adaptation (Hemmer-Hansen *et al.*, 2007; McKeown *et al.*, 2016, 2017) in systems where high gene flow would be expected to prevent such differentiation. In addition to resolving intraspecific population structure, genetic surveys have revealed substantial cryptic species diversity in the marine realm (Thorpe *et al.*, 2000; Borsa, 2002). The recognition of such distinct components is now regarded as fundamental to sustainable management.

The blue or spangled emperor fish, *Lethrinus nebulosus* (Forsskål, 1775), occurs throughout the Indo-west central Pacific (Fischer and Bianchi, 1984; Grandcourt *et al.*, 2006). Throughout its distribution, *L. nebulosus* is harvested by various artisanal, recreation, and commercial fisheries (Fischer and Bianchi, 1984; Heileman *et al.*, 2015). Although historically a minor component of some regional fisheries, the species is being exposed to increased fishing pressure as catches of other species decline

(Mann, 2000). Many *L. nebulosus* fisheries are already heavily exploited, operating above optimal sustainable levels and showing evidence of both recruitment and growth overfishing (e.g. Grandcourt *et al.*, 2006). Fishing pressure, coupled with habitat loss and alteration, is proposed to have led to declines in the biomass and abundance (or the complete absence) of *L. nebulosus* from certain areas, as well as declines in yields and catch per unit effort across its distribution (Heileman *et al.*, 2015). Overfishing can also affect ecological processes (Hutchinson and Reynolds, 2004) and drive genetic changes that can seriously compromise adaptability (Iles and Sinclair, 1982), resilience, and productivity (Ryman *et al.*, 1995).

The initial objective of this research was to describe population genetic variation of *L. nebulosus* in the Southwest Indian Ocean (SWIO), a region described as a biodiversity hotspot (Ridgway and Sampayo, 2005) and an area wherein fishing pressure on *L. nebulosus* is predicted to increase. Specifically, we aimed to test the null hypothesis of high connectivity throughout the region and compare patterns of genetic variation to other geographical regions to investigate if there were any signals of genetic erosion. However, nuclear and mitochondrial data revealed two highly divergent genetic groups indicating that, despite no *a priori* evidence, hitherto described *L. nebulosus* in the SWIO actually comprise two cryptic species that are being indiscriminately harvested.

Material and methods

Sample collection and DNA extraction

Sampling was coordinated by the SWIO Fisheries Project (van der Elst *et al.*, 2009). Samples (fin/muscle tissue/gillraker clips fixed in 95% ethanol) were collected from individuals identified as *L. nebulosus* based on morphology between 2011 and 2012, either on-board commercial fishing vessels or at commercial or subsistence landing sites across the SWIO (Figure 1). Total genomic DNA was extracted from samples using a combination of standard CTAB-chloroform/isoamyl alcohol methods (Winnepenninckx *et al.*, 1993) alongside Qiagen (Hilden, Germany) DNeasy Blood and Tissue and Sigma-Aldrich (St Louis, Missouri) GenElute Mammalian Genomic DNA Miniprep kits.

For genetic analyses, samples obtained from multiple localities within a country [e.g. Seychelles, (southern) Mozambique, and South Africa] were pooled within each country. Samples from Madagascar were obtained from the east, northwest, and southwest coasts. Owing to the longitudinal range of the island, those from the east and southwest coast were considered independently, while those from the northwest coast were combined with samples from the Comoros (owing to the small samples sizes from each of the individual localities) and designated as a Northern Mozambique Channel regional sample (Figure 1, Supplementary Table S1).

Control region and cytochrome oxidase I amplification and analysis

A fragment of the mtDNA Control Region (CR) was amplified by polymerase chain reaction (PCR) using newly designed species-specific primers (LCR1F 5′-CGGTCTTGTAAACCGGATGT-3′ and LCR1R 5′-GTCATGGCCCTGAAATAGGA-3′). PCRs were performed in a total volume of 20 µl, containing 4 µl template DNA, 2 mM $MgCl_2$, 0.5 µM forward primer and 0.5 µM of reverse primer, 0.2 mM dNTP mix, 1× reaction buffer, and 5 U *Taq* polymerase (BIOTAQ). The PCR thermoprofile was 180 s at 95 °C, followed by 45 cycles of 30 s at 95 °C, 45 s annealing at 54 °C and 60 s at 72 °C, with a final extension period of 10 min at 72 °C. PCR products were then purified using EXOSAP-IT and sequenced in both directions on an Applied Biosystem 3500 platform using the respective PCR primers. Sequences were trimmed manually, and 421 bp aligned across individuals using BIOEDIT (Hall, 1999). Additionally to facilitate a more comprehensive BLAST analysis, a 443 bp region of the barcoding COI gene was amplified in a subset of individuals representative of CR clades (described in Results), using universal primers FishF1 and FishR1 (Ward *et al.*, 2005) and similar PCR condition as those for the CR, with annealing temperature reduced to 55 °C.

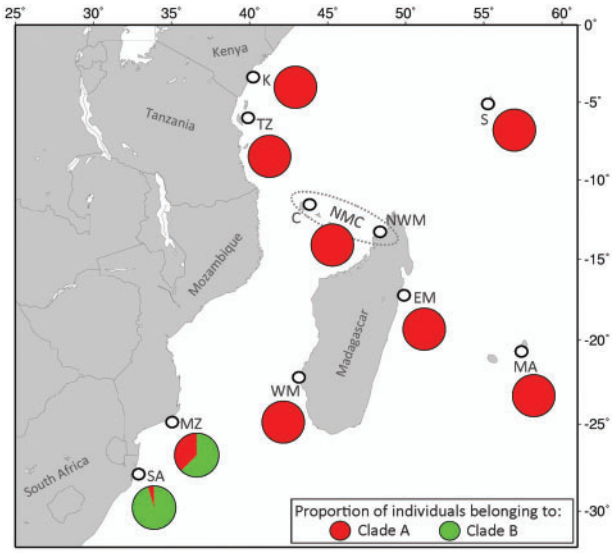

Figure 1. Sampling localities across the SWIO for L. nebulosus. Sampling locations are represented by circles and in some cases are midpoints of several geographically close sampling locations (see supplementary information for detailed sampling strategies). S = Seychelles, K= Kenya, TZ = Tanzania, C= Comoros, NWM= Northwest Madagascar, WM= West Madagascar, MZ = Mozambique, SA= South Africa, EM= East Madagascar, MA= Mauritius. Owing to small sample sizes C and NWM were grouped together and classed as NMC (Northern Mozambique Channel). Pie charts represent the proportion of individuals in each sampling site that belonging to mtDNA clade A and B.

Phylogenetic relationships among sequences were assessed using maximum likelihood (ML) and Bayesian inference (BI) trees, constructed in MEGA v7.0.21 (Tamura *et al.*, 2013) and MRBayes v3.2 (Ronquist and Huelsenbeck, 2003), respectively, using the optimum mutation model (HKY+G+I) inferred by JMODELtest (Posada, 2008). ML bootstrap (BS) values were calculated using 1000 BS replicates. BI phylogenies were calculated assuming unknown model parameters, and run over 5 000 000 generations, sampling the Markov chain every 1000 generations, with the first 15% of trees discarded as burn-in. Median joining haplotype networks were calculated and constructed in NETWORK (Bandelt *et al.*, 1999).

Genetic variation was described using number of haplotypes (H) in addition to indices of haplotype (h) and nucleotide (π) diversity alongside their variances, all of which were calculated in Arlequin v3.5.2.2 (Excoffier and Lischer, 2010). Genetic

differentiation among samples was assessed using global and pairwise Φ_{ST} values, implemented in Arlequin with significance tested after 10 000 permutations.

Microsatellite amplification and analysis

Nuclear genetic variation was assessed at 14 microsatellite loci developed for *L. miniatus* (Van Herwerden *et al.*, 2000a, 2000b), using protocols refined for *L. nebulosus* by Berry *et al.* (2012). Fragments were sized, alleles were scored, and a genotype matrix was produced using GeneMapper 5 (Applied Biosystems).

Micro-Checker 2.2.3 (Van Oosterhout *et al.*, 2004) was used to test for signatures of technical artefacts. The number of alleles (NA), allelic richness (AR), the absolute number of private alleles (N_P), the percentage of polymorphic loci ($P_{\%}$), and observed (H_O) and expected (H_E) heterozygosities were calculated using GenAlex 6.502 (Peakall and Smouse, 2012). Deviations from Hardy–Weinberg (HW) expectations across loci and samples were assessed using exact tests in GENEPOP 4.5.1 (Rousset, 2008), and linkage disequilibrium was examined using exact tests in Arlequin.

Genetic structure was investigated without prior sample definition, using the model-based Bayesian clustering programme STRUCTURE V2.3.4 (Pritchard *et al.*, 2000). Data were explored using combinations of admixture/no admixture models with 500 000 MCMC repetitions and a burn-in of 100 000 iterations. Twenty replicates were conducted for each number of clusters (K), and the K value that best fit the dataset was estimated using the log probability of data [Pr(X/K)] (Pritchard *et al.*, 2000). Clustering among individuals was also assessed using Discriminant Analysis of Principal Components (DAPC) implemented in ADEGENET (Jombart *et al.*, 2010). Additionally the relationship among samples based on allele frequencies was visualized using factorial correspondence analysis (FCA) in GENETIX (Belkhir *et al.*, 2004).

To quantify genetic differentiation, pairwise F_{ST} values (Wright, 1951) were calculated in Arlequin, with significance determined after 10 000 permutations. Hierarchical analysis of genetic variance (AMOVA; Excoffier *et al.*, 1992) was performed in ARLEQUIN to quantify, and test significance using 1000 permutations of, variation among groups (F_{CT}) and among samples within groups (F_{SC}), with groups defined by mtDNA clade. To further investigate the levels of nuclear differentiation between members of each mtDNA clade, "leave one out" self-classification tests were conducted in GENECLASS 2 (Piry *et al.*, 2004).

Results

mtDNA variability, phylogeny, and BLAST results

An edited 421-bp fragment of the CR was aligned across 292 individuals from 9 samples within the SWIO, producing a total of 234 haplotypes. Phylogenetic reconstruction (Figures 2 and 3) revealed the presence of two highly divergent clades within the SWIO, separated by an average of 50.341 nucleotide differences (SE = 6.001). Average sequence divergence between Clade A and Clade B was 11.96%. Pairwise comparisons between samples belonging to each clade revealed large and significant Φ_{ST} values in all comparisons ($\Phi_{ST} = 0.822$–0.889; Table 2). Overall haplotype diversity among Clade A individuals was 0.998 (±0.001) and Clade B was 0.991 (±0.004). Haplotype diversity for each clade was also high within each sample (Table 1).

Clade A was found across all SWIO samples, while Clade B was found exclusively in Mozambique and South Africa where it co-occurred with and was more abundant than Clade A. Using BLAST, Clade A CR haplotypes reported an 89% sequence similarity to *L. nebulosus* from New Caledonia (EU983053, EU983038, EU983064, EU983049, EU983040, EU983084, EU983075, EU983045, EU983030, EU983024, and EU983011), whereas clade B bore an 84% sequence similarity with *Lethrinus obsoletus* (AP009165) and *L. nebulosus* from New Caledonia and Bali (EU983030, EU983060, EU983027, EU983016, EU983090, EU983074, EU983040, EU983032, EU983020, and EU983051).

The separation of clades was also apparent among the representative COI sequences that revealed an average inter-clade sequence divergence of 5.64% (Supplementary Table S2). Analysis with BLAST of representative Clade A COI sequences reported 99% sequence similarity with voucher *L. nebulosus* sequences from Pomene in Mozambique (HQ561492 and JF493754) and formed a monophyletic cluster alongside *L. nebulosus* sequences from across the Indian Ocean and west Pacific (Supplementary Figure S1). Similarly, BLAST analysis of Clade B sequences reported 99% sequence similarity with voucher sequences of *L. nebulosus* from South Africa (DQ885022, DQ885021, DQ885020, and JF493753). BLAST analysis revealed that both clades were more similar to *Lethrinus ornatus* than to each other (Clade A: sequence divergence = 5.66%; Clade B: sequence divergence = 3.16%).

Microsatellite variability and congruence with mtDNA

A total of 242 individuals, across 8 samples (southwest Madagascar not genotyped), were genotyped at 14 microsatellite loci (Table 1). Microsatellite genotypes revealed clear partitioning of samples into two groups that aligned directly with clade membership. For example, the STRUCTURE analysis (Figure 4a) identified the most probable model to be $K = 2$ ($P = 1$ and 0 for all other K values tested), wherein members of both clades were separated with high probability (Q values > 0.8). STRUCTURE analysis also did not reveal any individuals with intermediate assignment probabilities that might indicate hybrids or admixed genotypes. GENECLASS assignment tests also revealed 100% classification of individuals (165 for Clade A and 77 for Clade B) according to clade. This nuclear/mtDNA disequilibrium was also apparent in the FCA plots (Figure 4b), which identified two distinct clusters of individuals corresponding to Clade A and B ancestry. Additionally, pairwise comparisons between samples partitioned according to mtDNA clade revealed high and significant F_{ST} values in all comparisons ($F_{ST} = 0.084$–0.115; Table 2). AMOVA of microsatellite genotypes partitioned according to mtDNA clade revealed a significant component of the nuclear variation (8.84%, $F_{CT} = 0.088$; $p < 0.001$) present was attributed to differentiation among the two mtDNA clades with much lower, but still significant structuring within clades (0.78% $F_{SC} = 0.009$, $p < 0.001$).

Within-clade structuring

As both mtDNA clades co-occurred among the Mozambique and South African samples, these individuals were partitioned by clade membership for subsequent tests of structure within each clade. For the microsatellite dataset, no scoring errors attributed to large allele drop-out or stuttering were detected; however, several significant instances of the possible presence of null alleles (11 of 140 loci by sample comparisons) and linkage disequilibrium (53 of 1007 loci by loci by sample comparisons) were

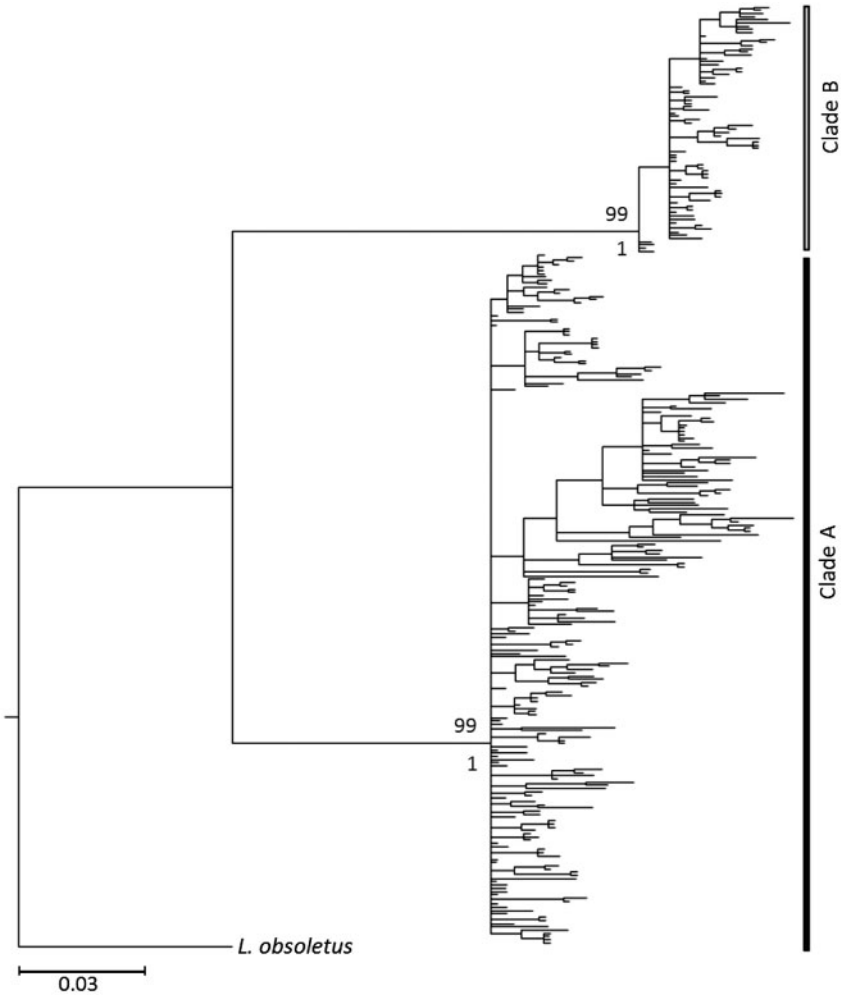

Figure 2. Phylogenetic relationships within SWIO *Lethrinus nebulosus* clades A and B, using 421bp CR mtDNA. Statistical support for nodes is given for both ML analyses (bootstrap support) above branches and Bayesian analyses (posterior probabilities) below branches. *Lethrinus obsoletus* is used as an outgroup.

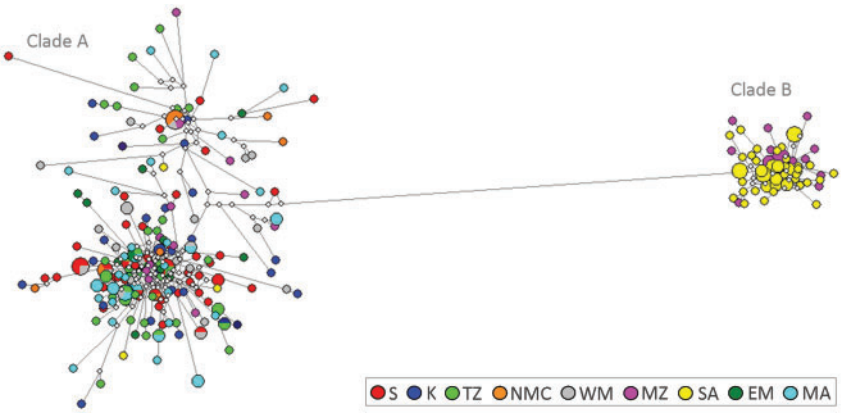

Figure 3. Median joining Haplotype network of L. nebulosus based on 421 bp CR mtDNA. Branch lengths are proportional to the number of differences. The node size is proportional to the haplotype frequency and node colour represents sampling location. Each small white circle represents a median vector. S = Seychelles, K= Kenya, TZ = Tanzania, NMC= Northern Mozambique channel, WM= West Madagascar, MZ = Mozambique, SA= South Africa, EM= East Madagascar, MA= Mauritius.

Table 1. Mitochondrial genetic diversity levels and neutrality tests for *L. nebulosus* based on (a) 421 bp of CR mtDNA and (b) microsatellite genetic diversity across 14 loci in *L. nebulosus*.

	mtDNA CR[a]				msats[b]							
	N	*H*	*h*	π	*N*	N_A	AR	$P_{\%}$	H_O	H_E	*P*	*F*
Seychelles	49	44	0.994 (0.006)	0.0370(0.0187)	36.143 ± 0.206	11.500 ± 2.062	3.637	100.00	0.699 ± 0.063	0.697 ± 0.064	0.258	0.000 ± 0.035
Kenya	29	28	0.998 (0.010)	0.0498(0.0252)	28.214 ± 0.459	10.143 ± 1.760	3.514	100.00	0.681 ± 0.060	0.682 ± 0.064	**<0.001**	−0.020 ± 0.034
Tanzania	41	37	0.995 (0.007)	0.0399(0.0202)	28.429 ± 0.644	9.929 ± 1.746	3.545	100.00	0.684 ± 0.051	0.696 ± 0.058	**<0.001**	−0.007 ± 0.042
NMC	10	7	0.867 (0.107)	0.0578(0.0315)	10.071 ± 0.339	6.500 ± 0.930	3.491	92.86	0.734 ± 0.075	0.653 ± 0.066	0.532	−0.130 ± 0.032
W. Madagascar	22	21	0.996 (0.015)	0.0492(0.0253)	–	–	–	–	–	–	–	–
Mozambique (Clades A & B)	32	27	0.984 (0.014)	0.1094(0.0542)	31.929 ± 0.071	12.714 ± 2.484	–	100.00	0.670 ± 0.071	0.714 ± 0.070	**0.002**	0.056 ± 0.033
Mozambique Clade A	12	12	1.000 (0.034)	0.0555(0.0297)	12.000 ± 0.000	7.714 ± 1.126	3.627	100.00	0.720 ± 0.071	0.681 ± 0.059	0.551	−0.052 ± 0.075
Mozambique Clade B	20	15	0.958 (0.033)	0.0188(0.0103)	19.929 ± 0.071	8.857 ± 1.667	3.484	85.71	0.639 ± 0.088	0.651 ± 0.083	0.357	0.025 ± 0.034
South Africa (Clades A & B)	60	45	0.989 (0.005)	0.0356(0.0179)	57.429 ± 0.644	13.214 ± 2.850	–	100.00	0.651 ± 0.076	0.678 ± 0.080	**<0.001**	0.029 ± 0.024
South Africa Clade A	3	3	1.000 (0.272)	0.0544(0.0417)	3.000 ± 0.000	3.214 ± 0.408	3.214	85.71	0.524 ± 0.125	0.536 ± 0.074	0.859	0.093 ± 0.155
South Africa Clade B	57	42	0.988 (0.006)	0.0199(0.0104)	54.429 ± 0.644	11.786 ± 2.553	3.515	100.00	0.658 ± 0.077	0.671 ± 0.080	**<0.001**	0.010 ± 0.023
E. Madagascar	13	13	1.00 (0.030)	0.0505(0.0269)	11.929 ± 0.071	7.857 ± 1.378	3.599	100.00	0.660 ± 0.074	0.658 ± 0.068	0.311	−0.008 ± 0.051
Mauritius	36	31	0.992 (0.008)	0.0471(0.0237)	29.286 ± 0.244	10.571 ± 2.027	3.634	92.86	0.683 ± 0.071	0.693 ± 0.071	**0.009**	0.007 ± 0.037
Total	292	234	0.998 (0.001)	0.0757(0.0025)	233.429 ± 2.104	19.643 ± 3.679	–	100.00	0.677 ± 0.063	0.728 ± 0.068	**<0.001**	0.066 ± 0.017

Statistically significant estimates ($p < 0.05$) are highlighted in bold. Instances where tests could not be performed are indicated as such (-). Standard deviation is given in parenthesis.

[a]mtDNA CR: mtDNA sample size = *N*, number of haplotypes = *H*, number of private haplotypes = *pHap*, haplotype diversity = *h*, and nucleotide diversity = π.

[b]msats: Mean sample size per locus = *N* (±standard error; SE), mean number of alleles per locus = N_A (±SE), allelic richness = AR, the percentage polymorphic loci = $P_{\%}$, mean observed heterozygosity per locus = H_O (±SE), mean expected heterozygosity per locus = H_E (±SE), probability of deviation from Hardy–Weinberg expectations = *P* and mean F_{IS} per locus = *F* (±SE).

detected. Fifteen of 140 loci by sample comparisons showed significant ($p < 0.05$) deviations from HW expectations. However, significant deviations from HW equilibrium and linkage disequilibrium were distributed randomly across loci and samples, and so all loci were included in subsequent analyses based on *a priori* sample definition.

Both clades exhibited some degree of intra-clade structuring that was more apparent at mtDNA than nuclear loci. Within Clade B, significant pairwise Φ_{ST} values (Table 2) were observed between Mozambique and South Africa in the CR dataset ($\Phi_{ST} = 0.319$), but corresponding microsatellite F_{ST} values were not significant ($F_{ST} = 0.006$). Within Clade A, significant pairwise Φ_{ST} values (Table 2) were observed between the most samples (excluding Seychelles–Tanzania, Seychelles–West Mozambique, Kenya–Mozambique, Kenya–South Africa, Tanzania–West Madagascar, Tanzania–Mauritius, West Madagascar–Mauritius, Mozambique–South Africa, and South Africa–North Mozambique Channel). However, pairwise comparisons of microsatellite genotypes identified broad-scale genetic homogeneity with the only significant F_{ST} values identified in comparisons between Mauritius and all SWIO samples, excluding east Madagascar and South Africa.

Discussion

The salient feature of this study was the concordant mtDNA and nuclear partitioning of samples into two distinct genetic groups. MtDNA revealed two reciprocally monophyletic clades. Average sequence divergence between clades (11.96%) far exceeded that within Clades A (2.95%) and B (1.09%), with this ratio well more than even conservative barcode gaps (Hebert *et al.*, 2004; Lefebure *et al.*, 2006). MtDNA data therefore indicate the occurrence of two highly divergent clades likely representing cryptic species. A common criticism of mtDNA-based taxonomy is that it represents a single, maternally inherited, locus (Collins and Cruickshank, 2013). However, analysis of nuclear microsatellites revealed robust partitioning of members of both clades with no evidence of hybrids, supporting a high level of bi-parental reproductive isolation.

There was a clear geographical pattern with Clade A found throughout all SWIO samples while Clade B was restricted to locations at the southwestern perimeter of the SWIO, where it co-occurred with Clade A but at a higher frequency (at least in the samples collected here). This pattern was also concordant with geo-referenced samples identified through BLAST, where Clade B sequences clustered with *L. nebulosus* sequences from the western coast of South Africa, whereas Clade A sequences showed greater similarity with samples from across the Indian Ocean and west Pacific. The co-occurrence of Clades A and B in southern Mozambique and northeast South Africa indicate a degree of spatial overlap in these clades. Bayesian clustering analysis indicated that this reflected mechanical mixing that was not accompanied by introgressive hybridization. This clearly highlights the potential for indiscriminate harvesting of both species at local levels. Similarly, another genetic study of SWIO lethrinids (Healey *et al.*, in review) identified the co-occurrence of a cryptic species among samples presumed to be entirely composed of, and being harvested as, *Lethrinus mahsena*. Such species misidentification may compromise sustainable management by leading to overestimation of stock abundances or inappropriate management measures based on incorrect biological parameters such as age at maturity (e.g. Griffiths and Heemstra, 1995). On a wider scale, it also contributes to a fundamental underrepresentation of species richness

Table 2. Pairwise estimates of genetic differentiation between *L. nebulosus* samples based on (a) Pairwise θ_{ST} values based on 421 bp mtDNA (below the diagonal) and (b) Pairwise F_{ST} values based on 14 microsatellite loci (above the diagonal).

		Clade A									Clade B	
		Seychelles	Kenya	Tanzania	NMC	W. Madagascar	Mozambique	South Africa	E. Madagascar	Mauritius	Mozambique	South Africa
Clade A	Seychelles	–	0.000	0.000	0.001		0.000	0.028	0.000	**0.013**	**0.085**	**0.090**
	Kenya	**0.461**	–	0.000	0.008		0.015	0.005	0.000	**0.012**	**0.107**	**0.109**
	Tanzania	0.007	**0.437**	–	0.009		0.012	0.012	0.004	**0.023**	**0.100**	**0.103**
	NMC	**0.350**	**0.214**	**0.317**	–		0.019	0.042	0.011	**0.022**	**0.092**	**0.104**
	W. Madagascar	0.008	**0.413**	0.012	**0.253**	–						
	Mozambique	**0.442**	0.029	**0.415**	**0.108**	**0.371**	–	0.038	0.013	**0.031**	**0.102**	**0.109**
	South Africa	**0.511**	0.010	**0.479**	0.150	**0.425**	0.071	–	0.023	0.017	**0.102**	**0.115**
	E. Madagascar	**0.099**	**0.302**	**0.081**	**0.157**	**0.067**	**0.247**	**0.310**	–	0.009	**0.084**	**0.090**
	Mauritius	**0.019**	**0.419**	0.003	**0.272**	0.010	**0.375**	**0.434**	**0.070**	–	**0.096**	**0.101**
Clade B	Mozambique	**0.852**	**0.822**	**0.846**	**0.843**	**0.839**	**0.837**	**0.889**	**0.847**	**0.827**	–	0.006
	South Africa	**0.868**	**0.853**	**0.865**	**0.870**	**0.865**	**0.867**	**0.885**	**0.871**	**0.853**	**0.319**	–

Statistical significance is indicated in bold.

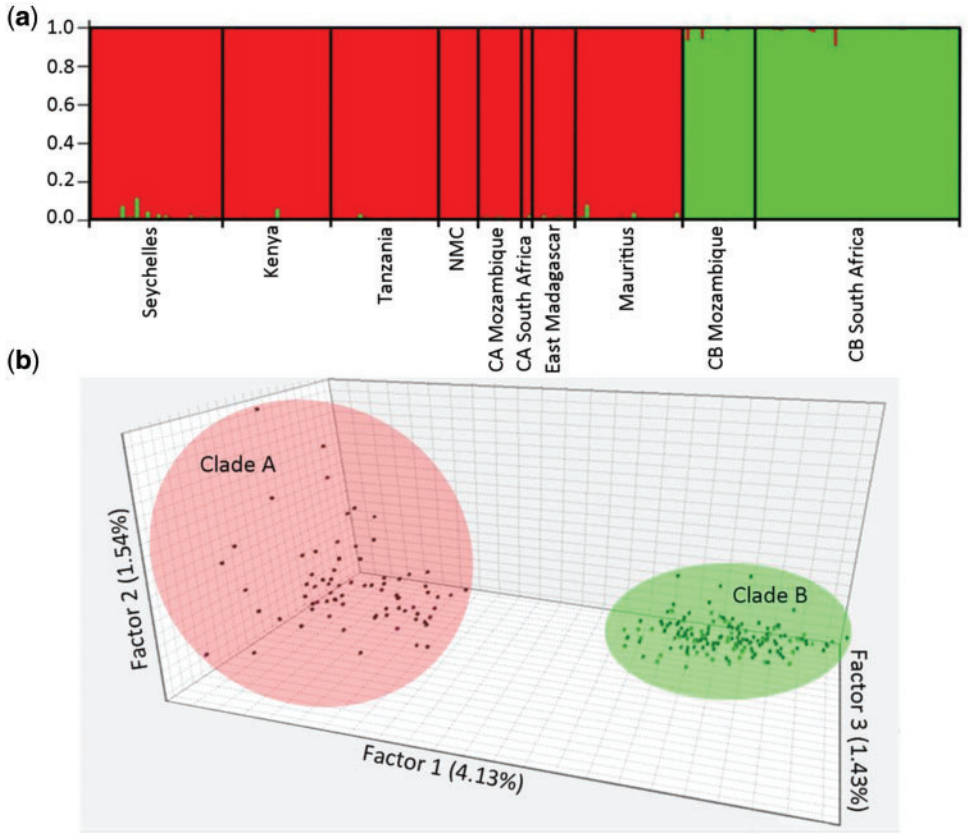

Figure 4. Estimates of genetic structure among 242 *Lethrinus nebulosus* individuals from eight regions in the south western Indian Ocean, based on genotypes from 14 polymorphic microsatellite loci. (a) STRUCTURE assignment of individuals to each of the two genetic groups (K = 2) identified. Each vertical line represents an individual with shading corresponding to the probability of that individual (or the proportion of the individual's genotypic ancestry) belonging to Population A or Population B. (b) Plot of individuals along three factors identified in the Factorial Correspondence Analyses, identifying two clusters corresponding to groups A and B.

and can cause inaccuracies in our understanding of biological, ecological, and evolutionary processes (Garcia-Vazquez *et al.*, 2012).

These data add to the number of genetic studies that have identified cryptic species among lethrinids (Borsa *et al.*, 2013; Healey *et al.* in review), and it has been suggested that the conservative phenotypes of the family may have contributed to such undetected diversity. In this regard, *L. nebulosus* is interesting as it exhibits a highly distinct phenotype relative to other lethrinids (Carpenter and Allen, 1989); hence, it is frequently referred to as the "spangled" emperor. Against the background of morphological similarity, the finding of two species with a similar, highly distinct phenotype that are not each other's closest phylogenetic

relative points to a mosaic of adaptive and plastic convergent effects.

For both species, microsatellites revealed limited evidence of structuring. Such broad-scale homogeneity is compatible with microsatellite patterns reported by Berry *et al.* (2012) for *L. nebulosus* in Australian waters. There was however some evidence of nuclear differentiation of Mauritian samples. Similar signals of isolation of Mauritian populations have been reported in other SWIO taxa, and have been attributed to a combination of large geographic distances from other coasts (>1000 km) and a barrier effect due to the landmass of Madagascar (e.g. Muths *et al.*, 2011, 2015).

In contrast to the nuclear diversity, mtDNA revealed a pattern of chaotic genetic patchiness, common among marine species (e.g. Selkoe *et al.*, 2006). This difference between marker types may reflect the greater susceptibility of mtDNA to genetic drift. In line with this, genetic patchiness has been variously attributed to episodic drift effects due to large variances in individual reproductive success (sweepstake recruitment) and/or larval cohesion. Sweepstakes recruitment has been reported for a number of highly fecund taxa and may generate genetic differentiation despite gene flow when recruitment is variable (McKeown *et al.*, 2017). Even in the absence of genetically isolated source populations, as might be the case here, larval cohesion (Selkoe *et al.*, 2006) may enhance and be effectively indistinguishable from sweepstake effects (Turner *et al.*, 2007). Berry *et al.* (2012), through analysis of age-segregated samples, empirically demonstrated the occurrence of larval cohesion in *L. nebulosus* and described how such cohesion and associated genetic differentiation were eroded by dispersal of sexually mature adults. As the samples analysed here are presumed to comprise multiple adults, the detection of such differentiation may indicate that samples are being harvested prior to sexual maturity, a practice likely to compromise sustainability. Similarly, patchiness among mixed adult samples must be considered a conservative reflection of any underlying recruitment variability (McKeown *et al.*, 2017).

This study has important implications for management and conservation of lethrinids in the SWIO. First, the resolution of two cryptic species within *L. nebulosus* highlights the inaccuracy of baseline biodiversity data for this group/region. In addition, the co-occurrence of both species within the southern samples suggests that both species are likely being indiscriminately harvested in this region. Such misidentification within fishery landings or stock assessments can severely compromise stock sustainability (Taylor *et al.*, 2012; McKeown *et al.*, 2015). Finally, stochastic recruitment suggested to underpin the observed genetic patchiness may decrease resilience of local stocks to fishing and increase unpredictability in recovery (Kuparinen *et al.*, 2014), and will necessitate a tailoring of the spatial scale of management (spatial bet hedging) according to biological and physical drivers of such recruitment variability. In addition, the detection of such patchiness may indicate that current fishing practises need to be amended to preclude harvesting of individuals that have not reached maturity.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Acknowledgements

This research was funded by the Southwest Indian Ocean Fisheries Project (SWIOFP) and through UK Natural Environment Research Council grants and was supported by the National Research Foundation (NRF) of South Africa and the South African Institute for Aquatic Biodiversity (SAIAB). We extend our gratitude to the SWIOFP Regional Management Unit for their support and assistance. We are grateful for the samples provided by the SWIOFP Observer Programme and the SWIOFP Component 3: Demersal Fisheries contact persons in the various countries. We thank Boaz Kaunda-Arara, Faratiana Ratsifa, Ibrahim Mohamed Toihir, Sabrena Lawrence, Vincent Lucas, Degambur Dharmendra, Rui Jorge Mutombene, and Teresa Athyade for their efforts in coordinating the sampling. We also thank Angus Macdonald (University of KwaZulu-Natal) for advice and assistance. Taryn Bodill and Elethu Duna are thanked for their assistance in the laboratory.

References

Bandelt, H. J., Forster, P., and Röhl, A. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution, 16: 37–48.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. 2004. GENETIX 4.05, Windows TM software for population genetics. Laboratoire génome, populations, interactions, CNRS UMR, 5000.

Berry, O., England, P., Marriott, R. J., Burridge, C. P., and Newman, S. J. 2012. Understanding age-specific dispersal in fishes through hydrodynamic modelling, genetic simulations and microsatellite DNA analysis. Molecular Ecology, 21: 2145–2159.

Borsa, P. 2002. Allozyme, mitochondrial-DNA, and morphometric variability indicate cryptic species of anchovy (Engraulis encrasicolus). Biological Journal of the Linnean Society, 75: 261–269.

Borsa, P., Hsiao, D. R., Carpenter, K. E., and Chen, W. J. 2013. Cranial morphometrics and mitochondrial DNA sequences distinguish cryptic species of the longface emperor (*Lethrinus olivaceus*), an emblematic fish of Indo-West Pacific coral reefs. Comptes Rendus Biologies, 336: 505–514.

Carpenter, K. E., and Allen, G. R. 1989. FAO Species Catalogue, 9. Food & Agriculture Organisation, Rome.

Collins, R. A., and Cruickshank, R. H. 2013. The seven deadly sins of DNA barcoding. Molecular Ecology Resources, 13: 969–975.

Excoffier, L., and Lischer, H. E. L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10: 564–567.

Excoffier, L., Smouse, P. E., and Quattro, J. M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics, 131: 479–491.

Fischer, W., and Bianchi, G. 1984. FAO Species Identification Sheets for Fishery Purposes: Western Indian Ocean (Fishing Area 51), II. Food and Agricultural Organization of the United Nations, Rome.

Garcia-Vazquez, E., Machado-Schiaffino, G., Campo, D., and Juanes, F. 2012. Species misidentification in mixed hake fisheries may lead to overexploitation and population bottlenecks. Fisheries Research, 114: 52–55.

Grandcourt, E. M., Al Abdessalaam, T. Z., Al Shamsi, A. T., and Francis, K. 2006. Biology and assessment of the painted sweetlips (*Diagramma pictum* (Thunberg, 1792)) and the spangled emperor (*Lethrinus nebulosus* (Forsskål, 1775)) in the southern Arabian Gulf. Fisheries Bulletin, 104: 75–88.

Griffiths, M. H., and Heemstra, P. C. 1995. A contribution to the taxonomy of the marine fish genus Argyrosomus (Perciformes: Sciaenidae), with descriptions of two new species from southern Africa. Ichthyological Bulletin, 65: 1–40.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *In* Nucleic Acids Symposium Series, 41, pp. 95–98. Information Retrieval Ltd., London, c1979–c2000.

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S., and Francis, C. M. 2004. Identification of birds through DNA barcodes. PloS Biology, 2: 1657–1663.

Heileman, S., Fennessy, S. T., and van der Elst, R. P. 2015. Demersal fisheries, a retrospective analysis of their status in the Southwest Indian Ocean. *In* Offshore Fisheries of the Southwest Indian Ocean: Their Status and the Impact on Vulnerable Species, pp. 213–285. Ed. by R. P. Van der Elst, and B. I. Everett. Oceanographic Research Institute, Special Publication, 10. 448pp.

Hemmer-Hansen, J., Nielsen, E. E., Frydenberg, J., and Loeschcke, V. 2007. Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus L.*). Heredity, 99: 592–600.

Hutchinson, J. A., and Reynolds, J. D. 2004. Marine fish population collapses: consequences for recovery and extinction risk. Bioscience, 54: 297–309.

Iles, T. D., and Sinclair, M. 1982. Atlantic herring: stock discreteness and abundance. Science, 215: 627–633.

Jombart, T., Devillard, S., and Balloux, F. 2010. Discriminant analysis of principal component: a new method for the analysis of genetically structure populations. BMC Genetics, 11: 94.

Knutsen, H., Olsen, E. M., Jorde, P. E., Espeland, S. H., André, C., and Stenseth, N. C. 2011. Are low but statistically significant levels of genetic differentiation in marine fishes 'biologically meaningful'? A case study of coastal Atlantic cod. Molecular Ecology, 20: 768–783.

Kuparinen, A., Keith, D. M., and Hutchings, J. A. 2014. Allee effect and the uncertainty of population recovery. Conservation Biology, 28: 790–798.

Lefebure, T., Douady, C. J., Gouy, M., and Gibiert, J. 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. Molecular Phylogenetics and Evolution, 40: 435–447.

Mann, B. Q. 2000. South African linefish status reports. Oceanographic Research Institute Special Publication, 7: 1–260.

McKeown, N. J., Arkhipkin, A. I., and Shaw, P. W. 2015. Integrating genetic and otolith microchemistry data to understand population structure in the Patagonian Hoki (Macruronus magellanicus). Fisheries Research, 164: 1–7.

McKeown, N. J., Arkhipkin, A. I., and Shaw, P. W. 2016. Regional genetic population structure and fine scale genetic cohesion in the Southern blue whiting Micromesistius australis. Fisheries Research, 185: 176–184.

McKeown, N. J., Hauser, L., and Shaw, P. W. 2017. Microsatellite genotyping of brown crab Cancer pagurus reveals fine scale selection and 'non-chaotic' genetic patchiness within a high gene flow system. Marine Ecology Progress Series, 566: 91–103.

McKeown, N. J., Robin, J. P., and Shaw, P. W. 2015. Species-specific PCR-RFLP for identification of early life history stages of squid and other applications to fisheries research. Fisheries Research, 167: 207–209.

Muths, D., Tessier, E., and Bourjea, J. 2015. Genetic structure of the reef grouper *Epinephelus merra* in the West Indian Ocean appears congruent with biogeographic and oceanographic boundaries. Marine Ecology-an Evolutionary Perspective, 36: 447–461.

Muths, D., Tessier, E., Gouws, G., Craig, M., Mwale, M., Mwaluma, J., Mwandya, A. et al. 2011. Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the SW Indian Ocean. Marine Ecology Progress Series, 443: 167–180.

Peakall, R., and Smouse, P. E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics, 28: 2537–2539.

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L., and Estoup, A. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity, 95: 536–539.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Molecular Biology and Evolution, 25: 1253–1256.

Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, 155: 945–959.

Ridgway, T., and Sampayo, E. M. 2005. Population genetic status of the Western Indian Ocean: what do we know? Western Indian Ocean Journal of Marine Science, 4: 1–10.

Ronquist, F., and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19: 1572–1574.

Rousset, F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources, 8: 103–106.

Ryman, N., Utter, F., and Laikre, L. 1995. Protection of intraspecific biodiversity of exploited fishes. Reviews in Fish Biology and Fisheries, 5: 417–446.

Selkoe, K. A., Gaines, S. D., Caselle, J. E., and Warner, R. R. 2006. Current shifts and kin aggregation explain genetic patchiness in fish recruits. Ecology, 87: 3082–3094.

Shaw, P. W., Turan, C., Wright, J. M., O'connell, M., and Carvalho, G. R. 1999. Microsatellite DNA analysis of population structure in Atlantic herring (*Clupea harengus*), with direct comparison to allozyme and mtDNA RFLP analyses. Heredity, 83: 490–499.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution, 30: 2725–2729.

Taylor, A. L., McKeown, N. J., and Shaw, P. W. 2012. Molecular identification of three co-occurring and easily misidentified octopus species using PCR-RFLP techniques. Conservation Genetics Resources, 4: 885–887.

Thorpe, J. P., Solé-Cava, A. M., and Watts, P. C. 2000. Exploited marine invertebrates: genetics and fisheries. Hydrobiologia, 420: 165–184.

Turner, T. F., Dowling, T. E., Marsh, P. C., Kesner, B. R., and Kelsen, A. T. 2007. Effective size, census size, and genetic monitoring of the endangered razorback sucker, *Xyrauchen texanus*. Conservation Genetics, 8: 417–425.

Van der Elst, R. P., Groeneveld, J. C., Baloi, A. P., Marsac, F., Katonda, K. I., Ruwa, R. K., and Lane, W. L. 2009. Nine nations, one ocean: a benchmark appraisal of the South Western Indian Ocean Fisheries Project (2008–2012). Ocean & Coastal Management, 52: 258–267.

Van Herwerden, L., Benzie, J., Peplow, L., and Davies, C. 2000a. Microsatellite Markers for Coral Trout (Plectropomus laevis) and Red Throat Emperor (Lethrinus miniatus). AIMS Report 32. Australian Institute of Marine Science, Townsville. 14 pp.

Van Herwerden, L., Benzie, J., Peplow, L., and Davies, C. 2000b. Microsatellite markers for coral trout (*Plectropomus laevis*) and red throat emperor (*Lethrinus miniatus*) and their utility in other species of reef fish. Molecular Ecology, 9: 1929–1931.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4: 535–538.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. 2005. DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 360: 1847–1857.

Winnepenninckx, B., Backeljau, T., and De Wachter, R. 1993. Complete small ribosomal-subunit RNA sequence of the chiton *Acanthopleura japonica* (Lischke, 1873) (Mollusca, Polyplacophora). Nucleic Acids Research, 21: 1670–1670.

Wright, S. 1951. The genetical structure of populations. Annals of Eugenics, 15: 323–354.

Handling editor: W. Stewart Grant