

Evolution / Évolution

# Cytochrome *b* barcoding, molecular systematics and geographic differentiation in rabbitfishes (Siganidae)

Sarah Lemer<sup>a,b</sup>, Didier Aurelle<sup>b</sup>, Laurent Vigliola<sup>c</sup>, Jean-Dominique Durand<sup>d</sup>,  
Philippe Borsa<sup>a,\*</sup>

<sup>a</sup> Institut de recherche pour le développement (IRD), UR 128, 98848 Nouméa, New Caledonia

<sup>b</sup> Dimar (UMR 6540 CNRS/université de la Méditerranée), station marine d'Endoume, rue de la Batterie-des-Lions,  
13007 Marseille, France

<sup>c</sup> Secretariat of the Pacific Community, BP D5, 98848 Nouméa cedex, New Caledonia

<sup>d</sup> Laboratoire 'Génome, populations, interactions, adaptation' (CNRS/IFREMER/UM2/IRD UR070),  
Station méditerranéenne de l'environnement littoral, 34200 Sète, France

Received 6 August 2006; accepted after revision 6 September 2006

Available online 19 October 2006

Presented by Jean Rosa

## Abstract

The fish genus *Siganus* (Siganidae) is widely distributed in the coastal habitats of all the tropical Indo-Pacific, with 28 nominal species recognized so far, based on general morphology and coloration patterns. A mitochondrial phylogeny of 16 Siganidae species, based on the partial nucleotide sequences of the cytochrome *b* gene, was produced. Individual haplotypes of given nominal species generally clustered at the extremity of long branches, thus validating the current taxonomy. However, *S. lineatus* haplotypes formed a paraphyletic group including *S. guttatus*, while *S. fuscescens* haplotypes were apparently splitted in two groups, calling for further investigation. *S. woodlandi* and *S. argenteus* formed a monophyletic group, as expected from their close morphological relatedness, although they were separated by a substantial, 14.5–16.3% nucleotide distance. Among eight species sampled from different locations across the Indo-West Pacific, *S. argenteus* and *S. spinus* showed the lowest degree of geographic differentiation, a result that correlated well with their extended pelagic larval stage. Fixation index estimates were high in all six other species tested (*S. doliatus*, *S. fuscescens*, *S. lineatus*, *S. puellus*, *S. punctatus*, *S. vulpinus*). The cytochrome *b* gene fragment chosen here proved useful as a barcode in Siganidae. **To cite this article:** *S. Lemer et al., C. R. Biologies 330 (2007).*

© 2006 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

## Résumé

**Identification par code-barre, systématique moléculaire et différenciation géographique chez les picots (Siganidae), sur la base des séquences nucléotidiques du gène du cytochrome *b*.** Les poissons du genre *Siganus* (Siganidae), caractéristiques des habitats côtiers de l'Indo-Pacifique tropical, comprennent 28 espèces nominales, définies sur la base de leur morphologie générale et de leurs patrons de coloration. Une phylogénie mitochondriale portant sur 16 espèces du genre a été produite à partir des séquences nucléotidiques d'un fragment du gène du cytochrome *b*. Les haplotypes d'une espèce donnée se regroupent généralement à l'extrémité de longues branches, validant de ce fait la taxinomie actuelle. Cependant, les haplotypes de *S. lineatus* forment un groupe paraphylétique avec *S. guttatus* et deux haplogroupes distincts sont observés chez *S. fuscescens*, résultats qu'il conviendrait

\* Corresponding author.

E-mail address: [Philippe.Borsa@noumea.ird.nc](mailto:Philippe.Borsa@noumea.ird.nc) (P. Borsa).

d'approfondir. *S. woodlandi* et *S. argenteus* forment un groupe monophylétique, comme cela avait été inféré du fait de leur ressemblance morphologique, quoique ces deux espèces soient séparées par une distance nucléotidique relativement élevée (14,5–16,3%). Des huit espèces échantillonnées en différentes localités de la région Indo-Ouest Pacifique, *S. argenteus* et *S. spinus* montrent le plus faible degré de différenciation géographique. Ce résultat s'avère concorder avec la longue durée de leur phase larvaire pélagique. Les valeurs estimées de l'indice de fixation sont élevées chez les six autres espèces testées (*S. doliatus*, *S. fuscescens*, *S. lineatus*, *S. puellus*, *S. punctatus*, *S. vulpinus*). Le fragment du gène du cytochrome *b* utilisé pour la présente étude s'avère être un bon marqueur pour l'identification à l'espèce (ou code-barre) chez les Siganidae. **Pour citer cet article :** *S. Lemer et al., C. R. Biologies 330 (2007).*

© 2006 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

**Keywords:** Rabbitfishes; Siganidae; Barcode; Indo-Pacific

**Mots-clés :** Picots ; Siganidae ; Code-barre ; Indo-Pacifique

## 1. Introduction

Good systematics and taxonomy are desirable for understanding the biogeographical patterns of species diversity. Of all the marine living organisms of the Indo-West Pacific, Teleostean fish probably are among the best-described, even though their systematics and taxonomy still need considerable research effort [1,2].

The fish family Siganidae ('rabbitfishes') comprises a single genus, *Siganus*, with 28 nominal species validated so far [3,4]. Species richness in Siganidae decreases from the central Indo-West Pacific region, where a majority of the species are thought to originate from, to the periphery [3,5]. This pattern has been observed for other groups of marine organisms [6]. Therefore, Siganidae constitute an interesting model-group to study speciation processes in the Indo-West Pacific, but their taxonomy first requires a precise re-evaluation. Remarkable homogeneity in those characters that usually allow one to diagnose species, such as numbers of fin spines and soft rays, tooth count and tooth shape and size, and broad across-species overlap in other characters, such as gill-raker and lateral-line scale counts [3,4] have earlier led to taxonomic confusion in Siganidae. The present taxonomy of Siganidae, based on Woodland's (1990) extensive revision [3] relies on colour patterns and gross morphological features, such as body proportions and relative length of snout [3,4]. Apart from an allozyme survey describing the genetic relationships among 10 species of the family, 9 of which were sampled from a localised region of the tropical West-Pacific [7], little emphasis has been placed so far on validating the current classification using genetic markers.

Fish of the family Siganidae occur in the coastal waters of the tropical Indo-Pacific. Two species occurring in the Red Sea, *S. luridus* and *S. rivulatus*, have recently

expanded their area of distribution into the Mediterranean across the Suez Canal [8]. The geographic distributions of rabbitfish range from the entire tropical Indo-West Pacific except the easternmost Pacific islands to regional, presumably, in part, as a consequence of differences in the dispersal potential of juveniles [3]. A proportion of species are originally endemic to restricted geographic areas such as the enclosed Red Sea (*S. rivulatus*) and Andaman Sea (*S. magnificus*), and isolated groups of islands or reefs (Maldives: *S. puelloides*; scattered reefs off Western Australia: *S. trispilos*; Fiji: *S. uspi*; Tonga: *S. niger*) [3]. One species (*S. unimaculatus*) has an antitropical distribution [3]. There might be a second example of antitropical distribution in *Siganus*: *S. woodlandi* [4] has so far been reported from New Caledonia only, but the *sedaka-hana-aigo*, an undescribed rabbitfish from Japan "that appears to be either *S. woodlandi* or a very close relative" ([4], and references therein) might be its antitropical counterpart. Pairs of presumed sibling species have been reported [3]. The latter have an essentially Indian/West Pacific distribution with some overlap in the central Indo-Malay region [3]. Among these, morphologically intermediate individuals thought to be hybrids between *S. doliatus* and *S. virgatus* have been reported [3,9], and hybrids between *S. guttatus* and *S. lineatus* are similarly expected [9].

Here, we analysed the nucleotide sequences of a fragment of the cytochrome *b* gene in 16 Siganidae species sampled across the Indo-West Pacific. Our objectives were: (1) to evaluate the potential of the cytochrome *b* gene as a species-specific marker (a DNA-barcode) in Siganidae; (2) to quantify genetic differences between sibling species with parapatric distribution; (3) to estimate the degree of genetic differentiation among populations within a few species differing by the width of their geographic distribution.

Table 1  
Samples of Siganidae and Acanthuridae analysed for the present survey

Species, Sampling location	Sampling date	N	Collector
<i>Acanthurus blochii</i> Southern lagoon, New Caledonia	Apr. 2006	2	P. Borsa
<i>Siganus argenteus</i> Southern lagoon, New Caledonia	June–July 2005; Mar. 2006	24	P. Borsa, J.-L. Justine, P. Villard
Ouvea lagoon, Loyalty Islands	Feb. 2006	24	M. Leopold
Truk, Caroline Islands, FSM	Apr. 2006	23	L. Vigliola
<i>S. corallinus</i> Southern lagoon, New Caledonia	June–July 2005	2	P. Borsa
<i>S. dohiatus</i> Southern lagoon, New Caledonia	June 2005–Apr. 2006	9	P. Borsa
Truk, Caroline Islands, FSM	Apr. 2006	4	L. Vigliola
<i>S. fuscescens</i> Southern lagoon, New Caledonia	June 2005–Feb. 2006	10	P. Borsa, J.-L. Justine
Southern lagoon, New Caledonia	2005	3	C. Mellin
Philippines	< May 2002	4	[10]
<i>S. guttatus</i> Makassar, Sulawesi, Indonesia	Sep. 2005	11	P. Borsa
<i>S. javus</i> Makassar, Sulawesi, Indonesia	Sep. 2005	4	P. Borsa
<i>S. lineatus</i> New Caledonia	Apr. 2002–Mar. 2006	13	P. Borsa
Makassar, Sulawesi, Indonesia	Sep. 2005	5	P. Borsa
Addu Atoll, Maldives	Aug. 2005	1	R.C. Anderson
Philippines	< May 2002	1	[10] <sup>1</sup>
<i>S. luridus</i> Lattaqiah, Syria	2002	7	M. Hassan
<i>S. puellus</i> Southern lagoon, New Caledonia	June–July 2005	12	P. Borsa
Truk, Caroline Islands, FSM	Apr. 2006	9	L. Vigliola
<i>S. punctatus</i> Southern lagoon, New Caledonia	June 2005; Mar. 2006	21	P. Borsa, J.-L. Justine
Makassar, Sulawesi, Indonesia	Sep. 2005	10	P. Borsa
Ouvea lagoon, Loyalty Islands	Feb. 2006	3	M. Leopold
Truk, Caroline Islands, FSM	Apr. 2006	12	L. Vigliola
Philippines	< May 2002	1	[10] <sup>2</sup>
<i>S. rivulatus</i> Lattaqiah, Syria	2002	7	M. Hassan
<i>S. spinus</i> Southern lagoon, New Caledonia	Jan. 2003; June 2005	1	P. Borsa
Moorea, French Polynesia	Apr. 2006	5	D. Lecchini
<i>S. vermiculatus</i> Makassar, Sulawesi, Indonesia	Sep. 2005	2	P. Borsa
<i>S. virgatus</i> Philippines	< May 2002	3	[10]
<i>S. vulpinus</i> Southern lagoon, New Caledonia	July 2005	9	P. Borsa
Truk, Caroline Islands, FSM	Apr. 2006	22	L. Vigliola
Philippines	< May 2002	1	[10]
<i>S. woodlandi</i> Southern lagoon, New Caledonia	June–July 2005	9	P. Borsa
Unidentified larvae Southern lagoon, New Caledonia	2004–2005	22	L. Carassou

N: sample size.

<sup>1</sup> Initially identified as *S. vermiculatus* (GENBANK AY190556; [10]) but barcode-identified as *S. lineatus* (present results);

<sup>2</sup> Initially identified as *S. punctatissimus* (GENBANK AY190543; [10]) but barcode-identified as *S. punctatus* (present results).

## 2. Materials and methods

Two hundred and sixty-two rabbitfish belonging to 15 nominal species, and 22 unidentified rabbitfish larvae were collected from seven locations in the Indo-West Pacific and the Mediterranean (Table 1). The fish were obtained from fishermen (all samples from Ouvéa, Makassar and Truk, and most samples from the southern lagoon of New Caledonia), or as side-samples of ecological surveys (the *S. spinus* sample of Moorea, and some samples from the southern lagoon of New Caledonia including all larvae, which were collected using light traps [11]). Voucher-specimens for all Siganidae species collected in the southern lagoon of New Caledonia and in Moorea were deposited at the Museum national d'histoire naturelle, Paris (MNHN 2006-1525 to 1537). Muscle tissue samples of *S. luridus* and *S. rivulatus*, preserved frozen at Station méditerranéenne de l'environnement littoral, Sète, were sent to us by F. Bonhomme. A piece of muscle tissue of a *S. lineatus* specimen from the Maldives, deposited at the Bishop Museum, Honolulu (BPBM 40390), was donated to us by J.E. Randall. For one more species (*S. virgatus*) we only had at our disposal cytochrome *b* gene sequences available from the GENBANK sequence database (<http://www.ncbi.nih.gov/Genbank>: accession nos. AY190546, AY190549, and AY190550). The samples from the Makassar Strait region (South Sulawesi, Indonesia) included a group of nine individuals whose coloration patterns ranged from typical *S. guttatus* to typical *S. lineatus* [3], through a range of intermediate patterns (for seven individuals), indicating possible introgression between the two species.

DNA was extracted from ethanol-preserved fin clips or muscle fragments using either the standard phenol–chloroform extractions with Phase-Lock Gel™ tubes (Eppendorf, Le Pecq, France), or the DNeasy® kit (Qiagen GmbH, Hilden, Germany). Individual DNA extracts were subjected to PCR amplification of a 326-bp fragment of the cytochrome *b* gene, using a pair of degenerate primers (forward: 5'-TCMGAYATCGCCACAGCC-TTCTC-3'; reverse: 5'-AGGAAGTGGAAKGC GAAG-AA-3') designed from the alignment of the cytochrome *b* gene sequences of Siganidae available in GENBANK (accession nos. AY190542 to AY190555). The amplification reactions were done in 96-well plates with each well containing 20  $\mu$ L reaction mixture (2  $\mu$ L template DNA, 2.5 mM MgCl<sub>2</sub>, 0.48 mM dNTP mix, 0.5  $\mu$ M of each primer and 0.5 *U* Taq polymerase (Promega, Madison WI, USA)) in a Robocycler™ thermocycler (Stratagene, La Jolla CA, USA). The PCR program consisted of 3 min DNA-denaturation (at 94 °C) fol-

lowed by 35 cycles of 1 min denaturation (94 °C)/1 min annealing (60 °C)/1 min elongation (72 °C). The PCR products were shipped to GATC Biotech (Konstanz, Germany), or Macrogen (Seoul, South Korea) for nucleotide sequencing and the sequences were subsequently retrieved from the company's Internet website.

All nucleotide sequences, which were deposited in GENBANK (accession numbers DQ898027 to DQ898094), were aligned, together with their homologues from the Philippines (GENBANK AY190543 to AY190556), using BIOEDIT [12]. The phylogenetic relationships among species in the genus *Siganus* were inferred from the matrix of haplotype sequences using three different algorithms: neighbour-joining (NJ [13]), maximum parsimony (MP [14]), and maximum likelihood (ML [15]). NJ and MP analyses were done using MEGA 3.1 [16]. ML analysis was done using Phylip, version 3.65 [17]. Nucleotide distances between haplotypes ( $\hat{d}$ ) were estimated according to Kimura's 2-parameter model of nucleotide substitution ([18], implemented in MEGA 3.1). Gaps were treated as deletions in pairwise comparisons of sequences (option 'Pairwise-Deletion' in MEGA 3.1). Haplotype frequency differences between populations were estimated using  $\hat{\theta}$ , Weir and Cockerham's fixation-index estimator, which apportions the genetic diversity among samples relative to the total diversity [19]. Tests of significance were done by comparing the observed  $\hat{\theta}$ -value with 1000 pseudo- $\hat{\theta}$  produced by random permutations of haplotypes among samples using GENETIX 4.05 [20]. Taking into account the information harboured in the nucleotide sequences of each allele along with allele frequencies, we used analysis of molecular variance (AMOVA [21]) implemented in ARLEQUIN 3.01 [22] to estimate  $\Phi_{st}$ , Excoffier et al.'s (1992) fixation index for nucleotide sequence data [21]. The  $\Phi_{st}$ -estimate produced by AMOVA is equivalent to a multiple-locus  $\hat{\theta}$  where each variable nucleotide site is considered as a different locus [23]. The significance of  $\Phi_{st}$ -estimates was tested using the non-parametric permutation algorithm implemented in ARLEQUIN 3.01.

## 3. Results

The mitochondrial phylogeny of Siganidae, based on a 300-bp fragment of the cytochrome *b* gene, displayed clusters of haplotypes or haplogroups atop long branches, generally supported by high bootstrap scores (Fig. 1). In all cases, a haplogroup corresponded to a nominal species, except for the pair *S. doliatus*/*S. virgatus*, whose haplotypes formed a single, tight cluster. The homogeneity of *S. fuscescens* was question-

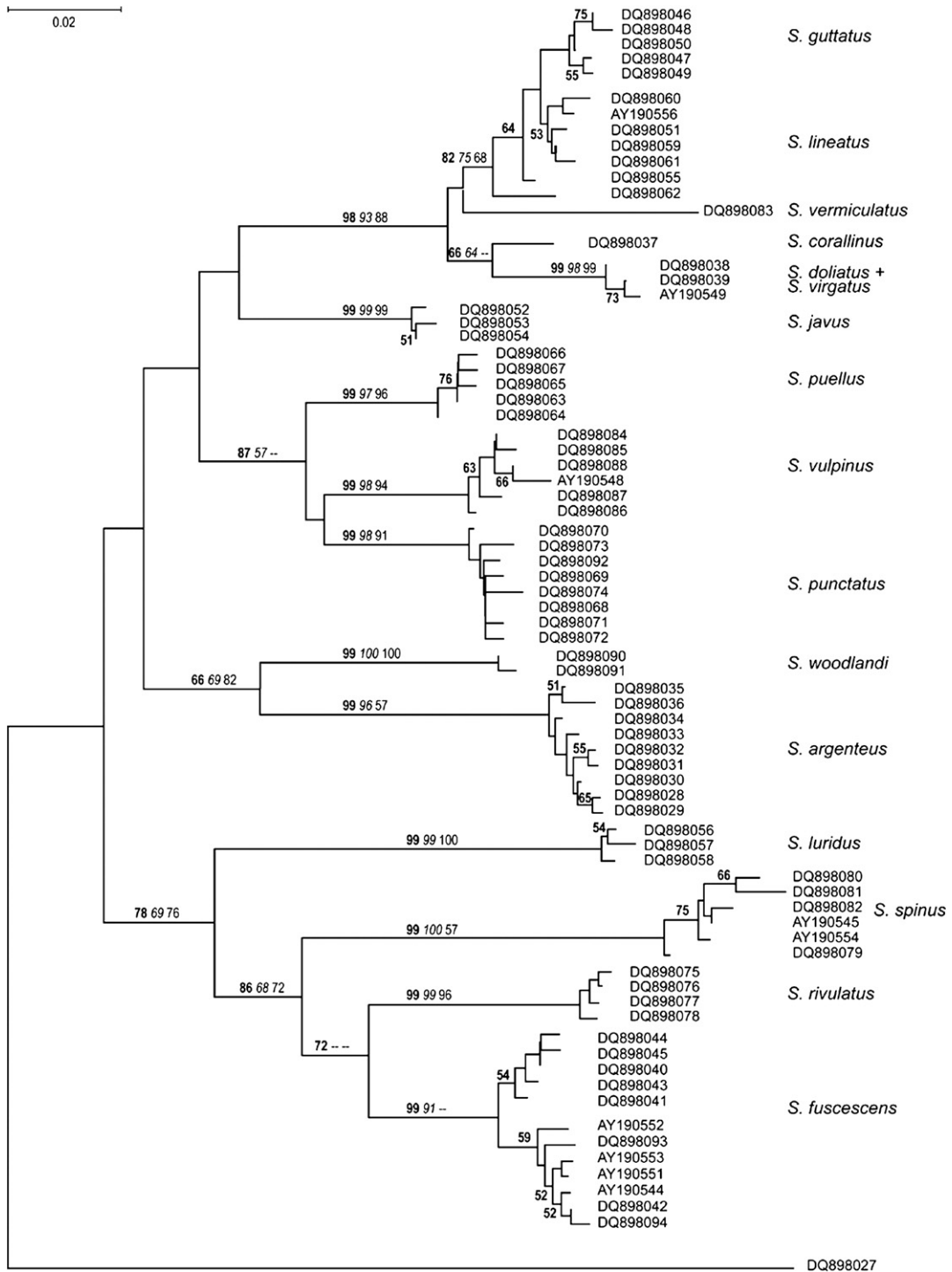


Fig. 1. Mitochondrial phylogeny of Siganidae (*Siganus* spp.). Neighbour-joining tree (Kimura's 2-parameter nucleotide distances [18]; MEGA 3.1 [16]) of nucleotide sequences of a 300-bp fragment of the cytochrome *b* gene (GENBANK DQ898028 to DQ898094 and AY190543 to AY190556). Homologous sequence in *Acanthurus blochii* (Acanthuridae) (GENBANK DQ898027) used as outgroup [24]. Numbers at a node are percentage scores (> 50%) after bootstrap resampling [25] of nucleotide sites in the matrix of sequences (bold: neighbour-joining, 1000 bootstraps; italics: maximum parsimony (MEGA 3.1), 1000 bootstraps; regular: maximum likelihood (PHYLIP 3.65 [17]), 100 bootstraps; for the sake of clarity, the two latter percentages were figured on the main nodes only). Scale bar: 2% nucleotide divergence estimate. Synonymous sequences: AY190549 = AY190550; DQ898039 = AY190546; DQ898047 = AY190547; DQ898054 = AY190555; DQ898068 = AY190543.

able, as the haplotypes sampled for that species appeared to form two separate haplogroups, distant by  $\hat{d} = 0.010\text{--}0.024$ . The haplogroup corresponding to *S. guttatus* was weakly supported and appeared to be nested within a *S. guttatus/S. lineatus* cluster, whose most external haplotype was that found in *S. lineatus* from the Maldives. The *S. guttatus* haplogroup was distant from the Indo-West Pacific (i.e., Makassar Strait, Philippines, New Caledonia) *S. lineatus* by  $\hat{d} = 0.010\text{--}0.020$  while the latter were separated from the Maldives *S. lineatus* haplotype by  $\hat{d} = 0.024\text{--}0.027$ . *S. luridus* and *S. spinus* were separated by a high nucleotide distance ( $\hat{d} = 0.145\text{--}0.163$ ) and the topology of the tree provided little support to the hypothesis that they form sister-clades. *S. vulpinus* haplotypes formed a sub-clade of a strongly supported clade that also included *S. puelus* and *S. punctatus*. *S. woodlandi* haplotypes formed a monophyletic clade with *S. argenteus*, distant from it by  $\hat{d} = 0.090\text{--}0.106$ .

The haplotypes of seven individuals morphologically intermediate between *S. guttatus* and *S. lineatus* were equally distributed among the two respective haplogroups characteristic of each species. The haplotypes of the 22 unidentified larvae sampled in the southern lagoon of New Caledonia clustered with either *S. lineatus* or *S. fuscescens*.

The degree of population differentiation, as estimated by fixation indices at the scale of the Indo-West Pacific, was significant in seven of the eight species tested (Table 2). The high degree of genetic differentiation found among *S. fuscescens* populations, however, could be misleading in the case where two reproductively isolated populations had been sampled in New Caledonia (see above). Accordingly, we re-estimated the fixation-index values after considering solely the *S. fuscescens* clade formed by haplotypes *DQ898042*, *DQ898093*, *AY190545*, and *AY190551* to *AY190553*, represented in both New Caledonia and the Philippines, and found  $\hat{\theta} = 0.250$  ( $P < 0.01$ ) and  $\hat{\phi}_{st} = 0.373$  ( $P < 0.01$ ). The fixation indices could not be estimated in the other *S. fuscescens* clade (formed by haplotypes *DQ898040*, *DQ898041*, *DQ898043* to *DQ898045*, and *DQ898094*) since the latter was represented in the New Caledonian sample only.

#### 4. Discussion

Individual haplotypes of given nominal species generally formed tight clusters at the extremity of long branches, thus validating the current, morphologically-based taxonomy. While *S. guttatus* haplotypes appeared to cluster into a single clade, thus confirming its recognition as a separate species (in spite of apparent hy-

bridization with *S. lineatus* (present results)), this group appeared to be nested within *S. lineatus*, suggesting that *S. guttatus* may be an offshoot of an ancestral proto-*S. lineatus* stock. The fact that the only haplotype sampled in the isolated *S. lineatus* population of the Maldives placed itself as outgroup to both *S. guttatus* and Indo-West Pacific *S. lineatus*, hints that the current taxon *S. lineatus* may include two geographically isolated populations with different evolutionary histories. Further insight into the genetic relationships of West-Pacific and Indian-Ocean populations of *S. lineatus*, requiring extensive sampling across the species' range is warranted, as this will inform us on the degree of genetic differentiation of Maldives *S. lineatus* vs. other *S. lineatus* populations. The two mitochondrial clades observed within *S. fuscescens* were separated from each other by an average nucleotide distance of the same order as that between *S. guttatus* and Indo-West Pacific *S. lineatus*, hinting at the possible occurrence of two distinct forms within the current taxon *S. fuscescens* in New Caledonia. To test that hypothesis requires further sampling of individuals, and the use of Mendelian markers to test for reproductive isolation.

Three pairs of presumed, sibling species with parapatric distribution were sampled for the present survey. Contrasted patterns of pairwise genetic differentiation were observed. The hypothesis of sibling species for the pair *S. luridus/S. spinus* was not supported. Letting aside the case of Maldives *S. lineatus*, the phylogenetic arrangement of *S. guttatus/S. lineatus* haplotypes was not inconsistent with the sibling-species hypothesis, but the fact that the two haplogroups did not appear as clearly separate clades also leaves open the possibility that the two species have not fully achieved reciprocal monophyly of their mitochondrial lineages. Nucleotide sequencing of a larger segment of the mitochondrial DNA may provide the necessary information to address that question. Last, there was no evidence of genetic differentiation of *S. doliatus* vs. *S. virgatus*. *S. virgatus* haplotypes were from individuals sampled in the Philippines, which border the zone of overlap with *S. doliatus* [3]. Since the two species are believed to hybridize [3,9], it is possible that those individuals identified as *S. virgatus* [10] harboured *S. doliatus* mitochondria as a result of introgression. In that case, further sampling of *S. virgatus* would be necessary, ideally outside the zone of overlap to sample 'true' *S. virgatus* mitochondria. Alternatively, the divergence between the two species may be recent, to the point that the sorting of their mitochondrial lineages is still not perceptible.

At the infra-specific scale, the degree of genetic differentiation was found to be lowest in *S. argenteus* and

Table 2

*Siganus* spp. mitochondrial haplotype (partial sequence of cytochrome *b* gene) frequencies by sample. All sequence data were from the present survey, except for 10 Philippines specimens, whose sequences were retrieved from GENBANK (details in footnote). Degree of genetic differentiation estimated using Weir and Cockerham's  $\hat{\theta}$  [19,20] and Excoffier et al.'s  $\hat{\phi}_{st}$  [21,22]. Nucleotide sites numbered from 1 to 300, with the first site homologous to site no. 14611 in Atlantic salmon, *Salmo salar* (GENBANK AF133701)

Species, Haplotype	Polymorphic site	Sample					$\hat{\theta}$	$\hat{\phi}_{st}$
		NC	Ouvéa	Truk	MS	Moorea		
	1222							
	35792479						0.020*	0.031
<i>S. argenteus</i>	61831931							
DQ898028	CACTGCCT	12	14	6	–	–		
DQ898029	.....C	0	0	1	–	–		
DQ898030	.....C	4	6	13	–	–		
DQ898031	.....TC	0	1	0	–	–		
DQ898032	...A.TC	2	0	0	–	–		
DQ898033	.G.....C	1	0	0	–	–		
DQ898034	.....A.C	0	1	0	–	–		
DQ898035	...C...C	4	1	3	–	–		
DQ898036	T.TC...C	1	1	0	–	–		
	12							
	55						0.352*	0.291*
<i>S. doliatus</i>	32							
DQ898038	TC	7	–	4	0	–		
DQ898039	.T	2	–	0	1 <sup>1</sup>	–		
AY190549	C.	0	–	0	2 <sup>1</sup>	–		
	11111222							
	2577925668277						0.152**	0.279*
<i>S. fuscescens</i>	37423339683636							
DQ898040	AACGTATTCAGGCG	3	–	–	0	–		
DQ898041	.....C.....	11	–	–	0	–		
DQ898042	.G.A.G.CT.....	11	–	–	0	–		
DQ898043	....C...T.....	1	–	–	0	–		
DQ898044	.....T.....	3	–	–	0	–		
DQ898045	.....A..	2	–	–	0	–		
DQ898093	GG...G.CT...A	1	–	–	0	–		
DQ898094	..A....T.....	1	–	–	0	–		
AY190544	.G.A.G.CT.A...	0	–	–	1 <sup>2</sup>	–		
AY190551	.G...G.CT.....	0	–	–	1 <sup>2</sup>	–		
AY190552	.G...C.CT...T.	0	–	–	1 <sup>2</sup>	–		
AY190553	.G...G.CTG....	0	–	–	1 <sup>2</sup>	–		
	11222							
	816379						0.131*	0.394***
<i>S. lineatus</i>	143634							
DQ898051	TCGTTCG	0	–	–	1	–		
DQ898055	CT...A	10	–	–	0	–		
DQ898059	CT....	3	–	–	4	–		
DQ898060	CT.GTA	1	–	–	0	–		
DQ898061	CTT...	1	–	–	0	–		
AY190556	CT.G..	0	–	–	1 <sup>3</sup>	–		
	1112							
	1882						0.101*	0.161*
<i>S. puellus</i>	7696							
DQ898063	GACG	11	–	4	–	–		
DQ898064	A...	1	–	0	–	–		
DQ898065	.G..	0	–	3	–	–		
DQ898066	..T.	0	–	1	–	–		
DQ898067	...A	0	–	1	–	–		

(continued on next page)

Table 2 (continued)

Species, Haplotype	Polymorphic site	Sample					$\hat{\theta}$	$\hat{\Phi}_{st}$
		NC	Ouvéa	Truk	MS	Moorea		
	1111222							
	122239789						0.053*	0.061
<i>S. punctatus</i>	511385527							
<i>DQ898068</i>	GTATTACTT	20	2	12	6 <sup>4</sup>	–		
<i>DQ898069</i>	.....G...	1	0	0	0	–		
<i>DQ898070</i>	...C.....	0	0	0	1	–		
<i>DQ898071</i>	.....A..	0	0	0	1	–		
<i>DQ898072</i>	.....C.	0	0	0	1	–		
<i>DQ898073</i>	A.....C	0	0	0	1	–		
<i>DQ898074</i>	.GT.....	0	1	0	0	–		
<i>DQ898092</i>	...G....	0	0	0	1	–		
	1122222							
	215233449						0.017	–0.214
<i>S. spinus</i>	713713030							
<i>DQ898079</i>	ACTTAATTC	1	–	–	0	0		
<i>DQ898080</i>	G..C...A	0	–	–	0	1		
<i>DQ898081</i>	G..CTGCCA	0	–	–	0	1		
<i>DQ898082</i>	GT.....	0	–	–	0	1		
<i>AY190545</i>	G.C.....	0	–	–	1	2		
<i>AY190554</i>	G.....	0	–	–	1	0		
	122222							
	25113699						0.105	0.316*
<i>S. vulpinus</i>	74767417							
<i>DQ898084</i>	ACACAATT	5	–	14	0	–		
<i>DQ898085</i>	.....G..	3	–	4	0	–		
<i>DQ898086</i>	.....C.	1	–	0	0	–		
<i>DQ898087</i>	G...G...	0	–	1	0	–		
<i>DQ898088</i>	.T.....	0	–	1	0	–		
<i>DQ898089</i>	.....C	0	–	2	0	–		
<i>AY190548</i>	.TGT....	0	–	0	1 <sup>5</sup>	–		

NC: southern lagoon of New Caledonia (for all species except *S. lineatus*, sampled on the western coast and in the northern lagoon); MS: Makassar strait off Sulawesi Island (for all samples except where noted);

\*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$ .

<sup>1</sup> Specimens from the Philippines identified as *S. virgatus* [10].

<sup>2</sup> From the Philippines [10].

<sup>3</sup> Specimen from the Philippines initially identified as *S. vermiculatus* (GENBANK AY190556; [10]) but barcode-identified as *S. lineatus* (present results).

<sup>4</sup> Includes one sequence of a Philippines specimen initially identified as *S. punctatissimus* (GENBANK AY190543; [10]) but barcode-identified as *S. punctatus* (present results).

<sup>5</sup> From the Philippines [10].

*S. spinus*, which of all Siganidae are the two species with the presumably longest pelagic, larval/prejuvenile stage [3] and also the widest distribution. Genetic-differentiation estimates in other species with more restricted ranges (*S. doliatus*, *S. fuscescens*, *S. lineatus*, *S. puellus*, *S. punctatus*, *S. vulpinus*) were all higher. Thus, distribution ranges, which primarily correlate with pelagic duration of larvae [3], also correlated well with the degree of genetic homogeneity among populations. These observations indirectly validate Woodland's (1990) hypothesis that dispersal is higher in species with longer larval stage in Siganidae [3].

The 300-bp cytochrome *b* gene fragment sequenced here proved a potential barcode for Siganidae, since

the haplotypes sampled in a species generally clustered tightly, relative to the long intervals between species. In other terms, intraspecific pairwise nucleotide distances were distinctly lower than pairwise distances between haplotypes of different species, allowing efficient discrimination. This molecular marker allowed us to identify *Siganus* spp. larvae, all of which the current, morphology-based identification keys can only assign to family [26]. Most of the cytochrome *b* gene sequences that were available on GENBANK in late 2005 when we started the present survey were from an unpublished study by D.G. Ochavillo and coauthors [10] whose objective was the identification and discrimination of juvenile Siganidae at the late pelagic presettlement stage.



Our results allowed the barcode identification of two of the individuals sequenced by these authors, that is AY190543, initially identified as *S. punctatissimus* but which was found to be *S. punctatus*, and AY190556, which was identified as *S. vermiculatus* but here found to be *S. lineatus*. Barcoding of Siganidae will be useful in our investigations on the population and community ecology of shorefish larvae and juveniles, which require identification at the species level [11,27,28].

## Acknowledgements

A part of this work was included in S. Lemer's M.Sci. memoir. Thanks to those who kindly provided samples: R.C. Anderson, F. Bonhomme, L. Carassou, J.-L. Justine, C. Mellin, G. Mou Tham, D. Lecchini, M. Leopold, J.E. Randall, P. Villard, and the Li-Khau family. Thanks to C. Rocher for technical help at the Dimar laboratory, and to P. Pruvost and L. Nandrin for curating voucher specimens at MNHN, Paris. Funded by the French government's 'Fonds Pacifique pour le développement', IRD-UR 128 'Biocomplexité des écosystèmes coralliens de l'Indo-Pacifique', Dimar, and ZoNéCo-Opération 'Jeunes poissons'.

## References

- [1] N. Knowlton, Molecular genetic analyses of species boundaries in the sea, *Hydrobiologia* 420 (2000) 73–90.
- [2] R.D. Ward, T.S. Zemlak, B.H. Innes, P.R. Last, P.D.N. Hebert, DNA barcoding Australia's fish species, *Philos. Trans. R. Soc. B* 360 (2005) 1847–1857.
- [3] D.J. Woodland, Revision of the fish family Siganidae with description of two new species and comments on distribution and biology, *Indo-Pac. Fishes* 19 (1990) 1–136.
- [4] J.E. Randall, M. Kulbicki, *Siganus woodlandi*, new species of rabbitfish (Siganidae) from New Caledonia, *Cybiurn* 29 (2005) 185–189.
- [5] D.J. Woodland, Zoogeography of the Siganidae: an interpretation of distribution and richness patterns, *Bull. Mar. Sci.* 33 (1983) 713–717.
- [6] S.R. Palumbi, What can molecular genetics contribute to marine biogeography? An urchin's tale, *J. Exp. Mar. Biol. Ecol.* 203 (1996) 75–92.
- [7] J.M. Lacson, S.G. Nelson, Genetic distances among fishes of the genus *Siganus* (Siganidae) from the western Pacific ocean, *Mar. Biol.* 116 (1993) 187–192.
- [8] A. Ben Tuvia, Two siganid fishes of Red Sea origin in the eastern Mediterranean, *Bull. Sea Fish. Res. Sta. Isr.* 37 (1964) 1–9.
- [9] J.E. Randall, Reef and Shore Fishes of the South Pacific, University of Hawai'i Press, Honolulu, 2005 (707 p.).
- [10] D.G. Ochavillo, P.M. Alino, G.J. Bakus, Identification and discrimination of late pelagic presettlement siganids: morphometric and mtDNA analyses, unpubl. <http://www.ncbi.nih.gov/Genbank>, 2002.
- [11] P. Borsa, L. Carassou, A. Collet, D. Ponton, Les premiers stades de vie des poissons de Nouvelle-Calédonie: diversité des espèces et rôle des zones côtières pour leur croissance et leur survie, Rapport pour Zonéco, IRD, Nouméa, 2005 (116 p.).
- [12] T.A. Hall, BIOEDIT: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, *Nucl. Acids Symp. Ser.* 41 (1999) 95–98.
- [13] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol.* 4 (1987) 406–425.
- [14] W.M. Fitch, Towards defining the course of evolution: minimum change for a specific tree topology, *Syst. Zool.* 20 (1971) 406–416.
- [15] J. Felsenstein, Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters, *Syst. Zool.* 22 (1973) 240–249.
- [16] S. Kumar, K. Tamura, M. Nei, MEGA 3: Integrated software for molecular evolutionary genetics analysis and sequence alignment, *Brief. Bioinf.* 5 (2004) 150–163.
- [17] J. Felsenstein, PHYLIP (Phylogeny Inference Package) Version 3.65, University of Washington, Seattle, 2005.
- [18] M. Kimura, A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences, *J. Mol. Evol.* 16 (1980) 111–120.
- [19] B.S. Weir, C.C. Cockerham, Estimating *F*-statistics for the analysis of population structure, *Evolution* 38 (1984) 1358–1370.
- [20] K. Belkhir, P. Borsa, L. Chikhi, N. Raufaste, F. Bonhomme, Genetix 4.05, logiciel sous Windows™ pour la génétique des populations, Laboratoire 'Génome, populations, interactions', CNRS UMR 5000, université Montpellier-2, Montpellier, France, 2004.
- [21] L. Excoffier, P.E. Smouse, J.-M. Quattro, Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data, *Genetics* 131 (1992) 479–491.
- [22] L. Excoffier, G. Laval, S. Schneider, ARLEQUIN ver. 3.0: An integrated software package for population genetics data analysis, *Evol. Bioinf. Online* 1 (2005) 47–50.
- [23] Y. Michalakis, L. Excoffier, A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci, *Genetics* 142 (1996) 1061–1064.
- [24] K.L. Tang, P.B. Berendzen, E.O. Wiley, G.D. Johnsson, R. Winterbottom, J.F. Morrissey, The phylogenetic relationships of the suborder Acanthuroidei (Teleostei: Perciformes) based on molecular and morphological evidence, *Mol. Phylogenet. Evol.* 11 (1999) 415–425.
- [25] J. Felsenstein, Confidence limits on phylogenies: an approach using the bootstrap, *Evolution* 39 (1985) 783–791.
- [26] J.M. Leis, B.M. Carson-Ewart, The Larvae of Indo-Pacific Coastal Fishes: An Identification Guide to Marine Fish Larvae, Brill, Leiden, The Netherlands, 2001 (850 p.).
- [27] L. Carassou, D. Ponton, Les premiers stades de vie des poissons de Nouvelle-Calédonie : identification des habitats lagonaires pour les stades pré- et post-installation, volet thématique 1 : Stades pré-installation, Rapport pour Zonéco, IRD, Nouméa, 2006 (56 p.).
- [28] P. Borsa, A. Collet, S. Lemer, Les premiers stades de vie des poissons de Nouvelle-Calédonie : identification des habitats lagonaires pour les stades pré- et post-installation, volet thématique 3 : Identification des jeunes stades, Rapport pour Zonéco, IRD, Nouméa, 2006 (23 p.).