# Mitochondrial DNA analysis of the geographic structure of Indian scad mackerel in the Indo-Malay archipelago

C. Perrin, Philippe Borsa

To be cited as:

PERRIN C., BORSA P. 2001. – Mitochondrial DNA analysis of the geographic structure of Indian scad mackerel, *Decapterus russelli* (Carangidae) in the Indo-Malay archipelago. *Journal of Fish Biology* 59, 1421-1426.

# Mitochondrial DNA analysis of the geographic structure of Indian scad mackerel, *Decapterus russelli* (Carangidae) in the Indo-Malay archipelago

C. PERRIN *†, P. BORSA *‡

* Laboratoire Génome Populations Interactions, CNRS UMR 5000, Station Méditerranéenne de l'Environnement Littoral, 1 Quai de la Daurade, 34200 Sète, France

† Present address: Department of Zoology, Otago University, P.O. Box 56, Dunedin, New Zealand

‡ Author to whom correspondence should be addressed. Tel. +33 4 67463390; fax +33 4 67463399; e-mail borsa@crit.univ-montp2.fr

Running title: Geographic structure of Indian scad mackerel

A fragment of the cytochrome *b* gene was analysed by PCR/SSCP in *Decapterus russelli* samples from the Indo-Malay archipelago. Sequence analysis revealed two distinct mitochondrial clades (mean nucleotide divergence=2.2%) whose geographic distribution was heterogeneous (Nei's $G_{ST}$=0.416), pointing to a complex pattern of genetic differentiation and demonstrating limited genetic exchange between populations in this highly mobile species.

Key words: cytochrome *b*; phylogeny; population genetics; Indo-West Pacific

Scad mackerels, *Decapterus* spp. are present in all tropical oceans. The Indian scad mackerel or layang biasa, *D. russelli* (Rüppell 1830), is one of the most common coastal pelagic species in South-East Asia where it is subjected to heavy exploitation by purse seiners (Widodo, 1988; Potier & Nurhakim, 1995; Sadhotomo, 1998). Hardenberg (1937) suggested that at least two stocks of layang are present in the Java sea and adjacent seas. One *D. russelli* stock may originate from the Flores sea and Makassar strait and reach the central-western part of the Java sea during the East monsoon; the other one originates in the Indian ocean and the Sunda strait and may reach the western part of the of the Java sea during the West (wet) monsoon, thus possibly being permanently separated from the former. Hardenberg (1937) also suggested the possible occurrence of a third layang stock occupying the southern part of the South China sea and reaching the northwestern part of the Java sea during the wet monsoon. In a preliminary survey of mitochondrial-DNA (control region) variation in *Decapterus* spp. based on restriction fragment length polymorphism (RFLP), Arnaud *et al.* (1999) observed a weak coalescence pattern among *D. russelli* haplotypes, all sampled in the Sulawesi sea. Neither polyphyly nor homoplasy in RFLP data could be discarded as explanations. Here we further analysed mitochondrial DNA haplotype frequencies and sequence phylogeny in *Decapterus russelli* from the Indo-Malay archipelago to tentatively (1) investigate the population genetics of this economically important but poorly known species; (2) test Hardenberg's hypothesis on stock structure; (3) test the species monophyly.

Three hundred and thirty six Indian scad mackerel were sampled in 1995 and 1997 from small or medium-sized purse seiners operating in the vicinity of 7 landing places in western Indonesia and Sabah (Fig. 1). Total genomic DNA was extracted using the phenol/chlorofrom protocol (Sambrook *et al.*, 1989) from alcohol-preserved muscle. A 355-base pair fragment of the mitochondrial DNA cytochrome *b* gene was amplified by polymerase chain reaction (PCR) using universal primers *CB1-L* (5'-CATCCAACATCTCAGCATGATGAAA-3') and *CB2-H* (5'-CCCTCAGAATGATATTTGTCCTCA-3') (Palumbi *et al.*, 1991). PCR amplifications were carried out in 25µl reaction mixture containing about 50-200 ng DNA, 0.8 m*M* dNTP mix, 2.5 m*M* $MgCl_2$, 1 × *Taq* buffer (Promega, Madison, USA), 0.5 µ*M* each primer and 0.25 unit *Taq* DNA polymerase (Promega). Thirty five cycles of PCR (denaturation at 94°C for 30 s, primer annealing at 53°C for 30 s, primer extension at 72°C for 1 min) were run in a Crocodile III thermocycler (Appligène, Strasbourg, France). Sequence polymorphism was detected by single-strand DNA conformation polymorphism (SSCP) analysis (Orita *et al.*, 1989). The single-strand DNAs were run in 1× TBE buffer at 4°C for 6 hrs at 15 W in non-denaturing polyacrylamide gel 1× MDE (FMC corporation, Rockland, USA). Gels were stained for 20 min in a solution of 1× TBE with 0.5 g/l ethidium bromide and destained for 5 min with deionised water, yielding single-strand DNA bands under ultraviolet light.

Sixteen different SSCP phenotypes were detected in the total sample. Phenotype *A* was the most common (frequency ≥0.90) in all samples except Toli-Toli. Sample Toli-Toli was dominated by phenotype *M*, which was absent from all the other samples (Fig. 2). Numerous other rare SSCP phenotypes were detected in the total sample. Ten individuals of SSCP phenotype *A*, 4 individuals of SSCP phenotype *M*, and all individuals characterised by rarer SSCP phenotypes were sequenced using the Thermosequenase sequencing kit of Amersham Life Sciences (Cleveland OH, USA) with either primer *CB1-L* or *CB2-H* on 5 µl of each PCR product purified for 15 min at 37°C by addition of 10 units of exonuclease I and 2 units of shrimp alkaline phosphatase (Amersham). There was no sequence variation among individuals of phenotype *A*, nor *M*. The two *bb* sequences from Labuan were identical, as were the three *e'* sequences from Labuan, Tambelan, and Kota Kinabalu. Phenotype *h*, which was found in 3 individuals from Labuan, corresponded to two sequences (*h1*, *h2*) differing from each other by one nucleotide transition. Phylogenetic inference on mtDNA sequence data was done using the computer package PHYLIP 3.57 (Felsenstein, 1995). Pairwise nucleotide distances between haplotypes were estimated using Kimura's (1980) 2-parameter model (procedure DNADIST of PHYLIP) with a ratio of two transitions to one transversion. Sequences of the homologous fragment in *Decapterus kurroides* Bleeker 1855, *D. macarellus* (Cuvier 1833), *D. macrosoma* Bleeker 1851, *D. tabl* Berry 1968, and *Selar crumenophthalmus* (Bloch 1793) were chosen as outgroups. The Neighbour-Joining phylogenetic tree (Fig. 2) generated from the matrix of nucleotide divergences (NEIGHBOR procedure of PHYLIP) and rooted by *S. crumenophthalmus* showed that all *D. russelli* mitochondria formed a monophyletic clade. Two phylogenetically distinct mitochondrial lineages (clades *A* and *M*), distant from each other by an average 2.2% nucleotide divergence (range: 1.3% to 3.0%), were found within *D. russelli*.

The presence of two distinct mitochondrial lineages (*A* and *M*) in Toli-Toli and two other samples (Fig. 2) is suggestive of past fragmentation followed by secondary contact, or of admixture in the samples of cryptic species. The amount of nucleotide divergence between *D. russelli* clades *A* and *M* suggests late-Pliocene or Pleistocene vicariance (Martin & Palumbi, 1993; Johns & Avise, 1998). While this may be much lower than the levels of divergence among other species in the genus *Decapterus* (Fig. 2), this was of the same magnitude as the lower divergence values at the cytochrome *b* locus among sister-species in other marine fishes (Johns & Avise, 1998) and between cryptic species of the genus *Beryx* (Hoarau & Borsa, 2000). It is therefore legitimate to address the question of whether *D. russelli* consists of a single species. Conversely, divergence values between mitochondrial clades similar or larger than that found between *D. russelli* clades *A* and *M* are not uncommon at the within-species level (e.g. Carr & Marshall, 1991; Baker *et al.*, 1995; Magoulas *et al.*, 1996; Miya & Nishida, 1997).

Phenotype frequency differences between populations were estimated using $G_{ST}$, Nei's (1973) parameter for the apportion of genetic diversity among samples relative to the total sample. Tests of significance were done by comparing the observed $G_{ST}$-value with 1000 pseudo-$G_{ST}$ calculated by random permutations of haplotypes using GENETIX (Belkhir *et al.*, 1996). The probability of the observation was estimated as $p=(n+1)/(N+1)$ where $n$ is the number of pseudo-$G_{ST}$ equal or greater than the observed $G_{ST}$ and $N$, the number of permutations (Sokal & Rohlf, 1996). Although the data failed to support Hardenberg's (1937) hypothesis of three stocks in the western part of the Indo-Malay archipelago (Java sea, South China sea, Indian ocean; $G_{ST}=0.033$; not significant), there was considerable geographical heterogeneity in the frequencies of the two haplogroups at the scale of the whole Indo-Malay archipelago (Fig. 2), with overall $G_{ST}=0.416$ ($p<0.001$). Pairwise comparisons of

samples designated the population sampled in Toli-Toli as the single outlier. Before discussing these results, it should be mentioned that they differ from those obtained for *Decapterus macrosoma* by Arnaud *et al.* (1999). Control-region RFLP haplotype frequencies did not vary significantly among *D. macrosoma* samples from the Sunda shelf and the Sulawesi sea, possibly reflecting the presence of a single population in this whole area, while a phylogeographic break was suspected in the Sunda strait (Arnaud *et al.*, 1999).

Under the single-species hypothesis, the geographic patterns of phenotype frequencies in *Decapterus russelli* would imply either geographic isolation, i.e. virtually no gene flow between Toli-Toli and the other populations, or selection. Invoking selection poses two problems: first, selection would assume that sharp habitat differences distinguish the southern Sulawesi sea from all the other areas sampled in the Indo-Malay archipelago. Second, an explanation should be sought for the fact that the Sulawesi sea would also be the only location where haplogroup polymorphism were virtually maintained. Polymorphism could be maintained by balance between selection and migration, but this would have to be unidirectional migration, from an *A* population to the Sulawesi sea population only. The idea that populations and species in the Indo-Malay archipelago may have undergone historical geographic isolation receives support from inferences on past sea levels, which left the Sunda shelf above water during colder periods in the Pleistocene (Tjia, 1980). This has been invoked to explain the present geographic structure of several marine species including butterflyfishes, *Chaetodon* spp. (McMillan & Palumbi, 1995), the giant tiger shrimp *Penaeus monodon* (Klinbunga, 1996), the starfish *Linckia laevigata* (Williams & Benzie, 1998), the mantis shrimp *Haptosquilla pulchella* (Barber *et al.*, 2000), and the clownfish *Amphiprion ocellaris* (Nelson *et al.*, 2000). Present-day geographical isolation may seem unlikely regarding the potentially high connectivity of populations of pelagic fishes such as *Decapterus russelli*, which travel large distances as eggs and larvae (Delsman, 1926), and adults (Hardenberg, 1937). In particular, oceanic flow through the Makassar strait potentially allows planktonic larvae from the Sulawesi sea to reach the southern Indonesian seas within weeks (Wyrtki, 1961; Barber *et al.*, 2000; Fig. 1). However, behaviour such as natal homing at the time of spawning may favour self-recruitment, hence explain the maintenance of haplotype-frequency differences between populations.

The hypothesis that clades *A* and *M* within *Decapterus russelli* actually represent the result of cryptic speciation (speciation without an obvious morphological signature) raises new questions. If the *A* and *M* mitochondrial clades respectively corresponded to say, species *A* and species *M*, then how to explain the geographic heterogeneity in haplotype frequencies within species *M*? Haplotype *M* would be fixed in the Sulawesi sea sample ($N$=20) while not being detected on the Sunda shelf (Pekalongan sample: $N$=1) nor in the Sunda strait (Labuan sample: $N$=3) which harboured distinct, private haplotypes of the *M* clade. This would again point to the geographic-isolation hypothesis. Alternatively, the boundary between cryptic species would coincide with geography (Sulawesi sea *vs.* the rest of the Indo-Malay archipelago) and the sorting of mitochondrial DNA lineages subsequent to speciation had not yet reached completion (i.e. reciprocal monophyly). This would imply an unlikely, complex history of speciation superimposed on secondary contact between historically separated populations. The additional use of nuclear-DNA markers is essential for testing the hypothesis of two cryptic species.

In conclusion, the cytochrome *b* gene sequence data do not favour the hypothesis of a polyphyletic *Decapterus russelli*, but instead indicate that control-region RFLP data (Arnaud *et al.*, 1999) were stained with homoplasy. The distinction of two clades within *D. russelli* is compatible with Pleistocene events that isolated the Sulawesi sea region

from other areas in the Indo-Malay archipelago. The SSCP phenotype-frequency data further indicate that presently there is limited effective genetic exchange between the Sulawesi sea and the other areas sampled, implying that Indian scad mackerel in the Indo-Malay archipelago should not be treated as a panmictic population. Since mitochondrial-DNA haplotype-frequency and sequence data do not suffice to choose among the several possible scenarios of genetic differentiation, nuclear-DNA markers should provide complementary insight.

We are grateful to P.H. Barber for helpful advice and corrections on this manuscript; to V. Castric, C. Poux, A. Rohfritsch, and Suprapto for help with laboratory work; to R. Andamari, S.B. Atmaja, A.S. Cabanban, G. Lesage, S. Nurhakim, M. Potier, Rumini, B. Sadhotomo, Suprapto, and Suwarso for help in collecting fish samples; to A.S. Cabanban and J.-M. Ecoutin for hospitality in Malaysia and Indonesia. We gratefully acknowledge the support of J.-R. Durand, J.-M. Ecoutin, M. Fatuchri Sukadi, P. Levang, S. Nurhakim, J. Roch, and J. Widodo (PELFISH: IRD and BPPL), F. Bonhomme, and J.-P. Féral. Funded by AB 632740 of IRD/DRV, IFREMER URM 16 and Réseau Diversité Marine.

## References

Arnaud, S., Bonhomme, F. & Borsa, P. (1999). Mitochondrial DNA analysis of the genetic relationships among populations of scad mackerel (*Decapterus macarellus*, *D. macrosoma* and *D. russelli*) in South-East Asia. *Marine Biology* **135**, 699-707.

Baker, C.S., Perry, A., Chambers, G.K. & Smith, P.J. (1995). Population variation in the mitochondrial cytochrome *b* gene of the orange roughy *Hoplostethus atlanticus* and the hoki *Macruronus novaezelandiae*. *Marine Biology* **122**, 503-509.

Barber, P.H., Palumbi, S.R., Erdmann, M.V. & Moosa, M.K. (2000). Biogeography: a marine Wallace's line? *Nature* **406**, 692-693.

Belkhir, K., Borsa, P., Goudet, J., Chikhi, L. & Bonhomme, F. (1996). *GENETIX version 3.0, Logiciel sous WINDOWS$^{TM}$ pour la Génétique des Populations*. Montpellier: Laboratoire Génome et Populations, Université Montpellier 2.

Carr, S.M. & Marshall, H.D. (1991). Detection of intraspecific DNA sequence variation in the mitochondrial cytochrome *b* gene of Atlantic cod (*Gadus morhua*) by the polymerase chain reaction. *Canadian Journal of Fisheries and Aquatic Sciences* **48**, 48-52.

Delsman, H.C. (1926). Fish eggs and larvae from the Java sea. 5. *Caranx curra*, *macrosoma*, and *crumenophthalmus*. *Treubia* **8**, 199-211.

Felsenstein, J. (1995). *PHYLIP - Phylogeny Inference Package, Version 3.57c*. Seattle: University of Washington.

Hardenberg, J.D.F. (1937). Preliminary report on a migration of fish in the Java Sea. *Treubia* **16**, 295-300.

Hoarau, G. & Borsa, P. (2000). Extensive gene flow within sibling species in the deep-sea fish *Beryx splendens*. *Comptes Rendus de l'Académie des Science de Paris, Série III* **323**, 315-325.

Johns, G.C. & Avise, J.C. (1998). A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. *Molecular Biology and Evolution* **15**, 1481-1490.

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111-120.

Klinbunga, S. (1996). *Genetic Variation and Population Structure of Giant Tiger Shrimp (*Penaeus monodon *Fabricius)*. PhD Thesis, Stirling: University of Stirling.

Magoulas, A., Tsimenides, N. & Zouros, E. (1996). Mitochondrial DNA phylogeny and the reconstruction of the population history of a species: the case of the European anchovy (*Engraulis encrasicolus*). *Molecular Biology and Evolution* **13**, 178-190.

Martin, A.P. & Palumbi, S.R. (1993). Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences of the USA* **90**, 4087-4091.

McMillan, W.O. & Palumbi, S.R. (1995). Concordant evolutionary patterns among Indo-West Pacific butterflyfishes. *Proceedings of the Royal Society of London, B* **260**, 229-236.

Miya, M. & Nishida, M. (1997). Speciation in the open ocean. *Nature* **389**, 803-804.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the USA* **70**, 3321-3323.

Nelson, J.S., Hoddell, R.J., Chou, L.M., Chan, W.K. & Wang, V.P.E. (2000). Phylogeographic structure of false clownfish, *Amphiprion ocellaris*, explained by sea-level changes on the Sunda shelf. *Marine Biology* **137**, 727-736.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. & Sekiya, T. (1989). Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Sciences of the USA* **86**, 2766-2770.

Page, D.M. (1996). An application to display phylogenetic trees on personal computers: TREEVIEW version 1.40. *Computer Applications in the Biosciences* **12**, 357-358.

Palumbi, S., Martin, A., Romano, S., McMillan, W.O., Stice, L. & Grabowski, G. (1991). *The simple fool's guide to PCR, version 2.0*. Honolulu: Department of Zoology and Kewalo Marine Laboratory, University of Hawaii.

Potier, M. & Nurhakim, S. (eds.) (1995). *Biodynex: Biology, Dynamics and Exploitation of the Small Pelagic Fishes in the Java Sea*. Jakarta, Indonesia: Indonesian Agency for agricultural Research and Development.

Sadhotomo, B. (1998). *Biologie des Principales Espèces Pélagiques Exploitées en Mer de Java*. Thèse de doctorat, Montpellier: Université Montpellier 2.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual, 2nd Edn*. New York: Cold Spring Harbor Laboratory.

Sokal, R.R. & Rohlf, F.J. (1996). *Biometry, the Principles and Practice of Statistics in Biological Research, 3rd Edition*. New York: Freeman and Co.

Tjia, H.D. (1980). The Sunda Shelf, southeast Asia. *Zeitschrifte für Geomorphologie* **24**, 405-427.

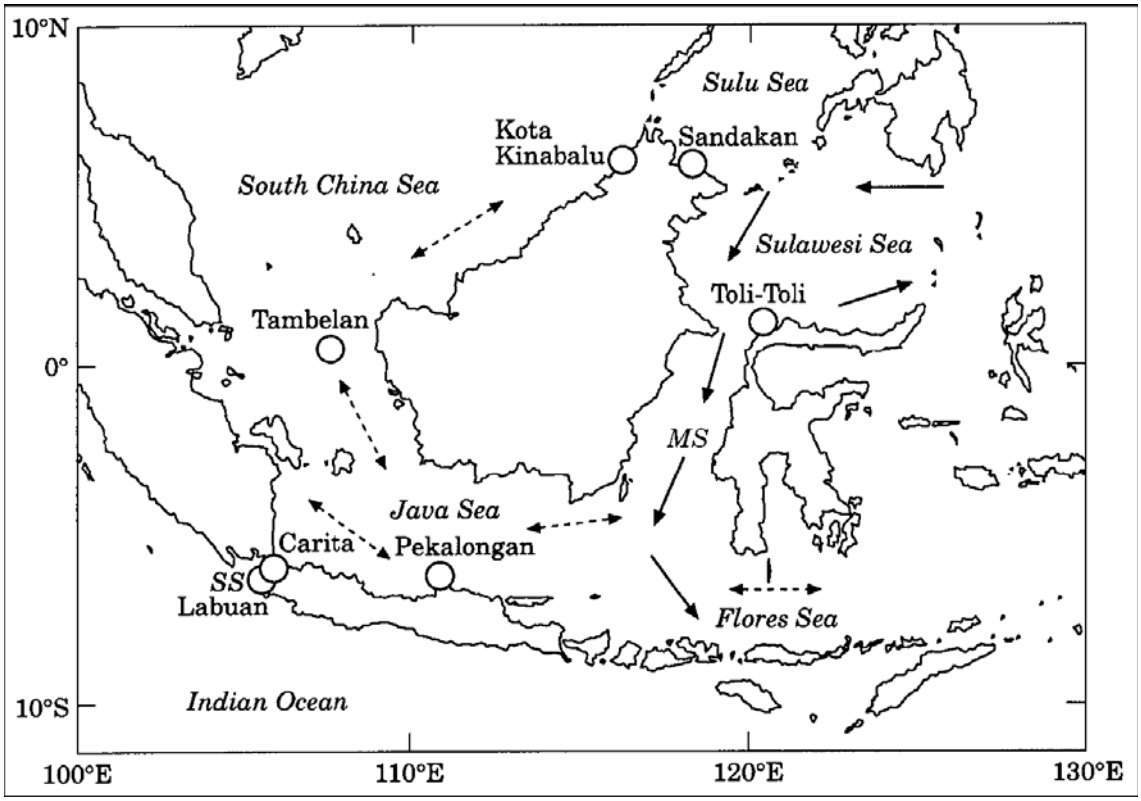
Widodo, J. (1988). *Population Dynamics and Management of "Ikan Layang", Scad Mackerel,* Decapterus *spp. (Pisces: Carangidae) in the Java Sea*. PhD dissertation, Seattle: University of Washington.

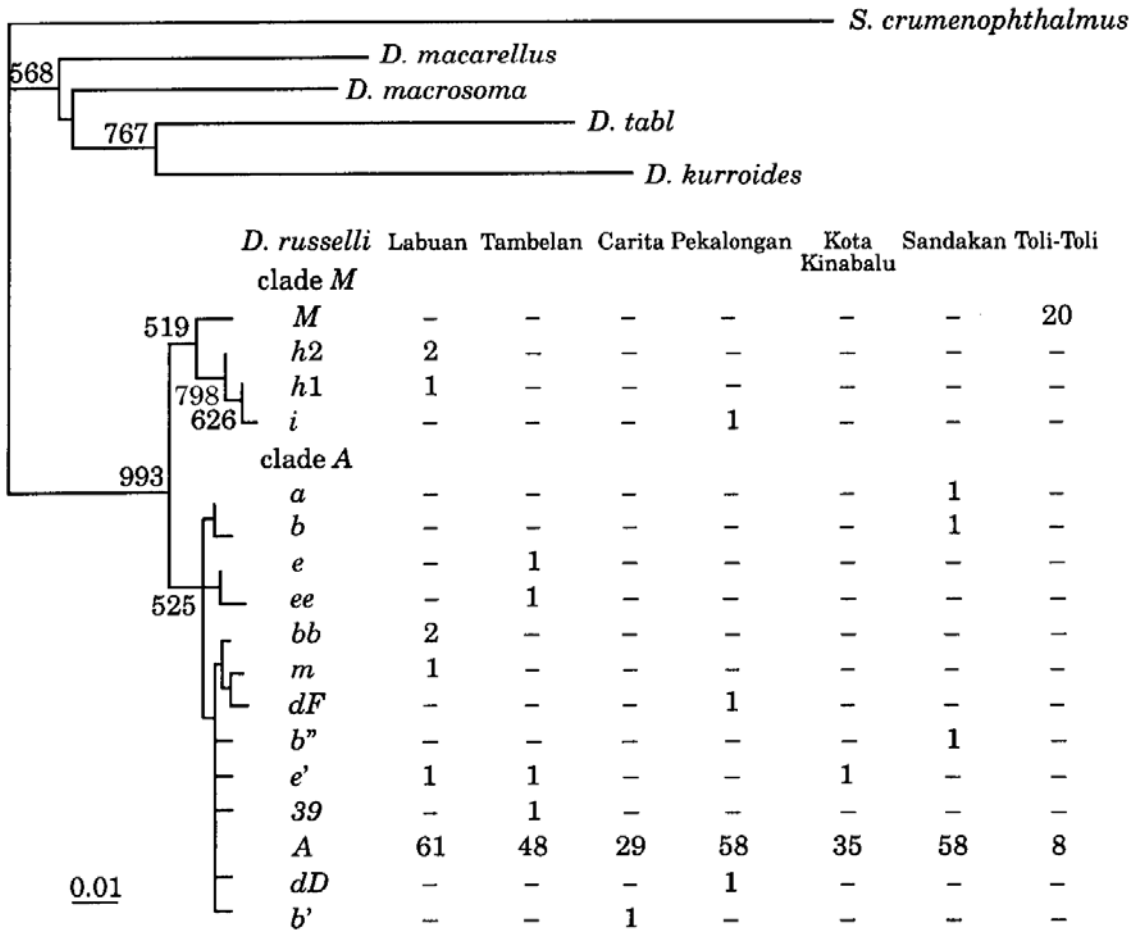
Williams, S.T. & Benzie, J.A.H. (1998). Evidence of a biogeographic break between populations of a high dispersal starfish: congruent regions within the Indo-West Pacific defined by color morphs, mtDNA, and allozyme data. *Evolution* **52**, 87-99.

Wyrtki, K. (1961). Physical oceanography of the southeast Asian waters. *Naga Report, Scripps Institute of Oceanography, University of California* **2**, 1-95.

FIG. 1. *Decapterus russelli*. Sampling locations (open circles) in the Indo-Malay archipelago. *MS*, Makassar strait; *SS*, Sunda strait. Bold arrows: dominant currents; dotted arrows: seasonal currents (Wyrtki, 1961).

FIG. 2. *Decapterus russelli*. Neighbour-joining tree [NEIGHBOR procedure in the PHYLIP package (Felsenstein, 1995); tree drawn under TREEVIEW (Page, 1996)] of partial nucleotide sequences (307 bp) of the cytochrome *b* gene for 17 haplotypes detected in 7 samples from the Indo-Malay archipelago, and inferred haplotype frequency by sample. The name of each sample (Fig. 1) has been abbreviated to its first four letters. All *D. russelli* sequences were deposited in GenBank (accession nos. AF307494-AF307510). Numbers at a node are scores out of 1000 bootstraps, obtained using procedure SEQBOOT of PHYLIP. Unlabelled nodes had bootstrap scores <500. Scale bar = 1% nucleotide divergence.


10°N
0°
10°S
100°E
110°E
120°E
130°E
Sulu Sea
Kota
Kinabalu
Sandakan
South China Sea
Sulawesi Sea
Toli-Toli
Tambelan
MS
Java Sea
Carita
Pekalongan
SS
Labuan
Flores Sea
Indian Ocean




| *D. russelli* | Labuan | Tambelan | Carita | Pekalongan | Kota Kinabalu | Sandakan | Toli-Toli |
|---|---|---|---|---|---|---|---|
| clade *M* | | | | | | | |
| *M* | – | – | – | – | – | – | 20 |
| *h2* | 2 | – | – | – | – | – | – |
| *h1* | 1 | – | – | – | – | – | – |
| *i* | – | – | – | 1 | – | – | – |
| clade *A* | | | | | | | |
| *a* | – | – | – | – | – | 1 | – |
| *b* | – | – | – | – | – | 1 | – |
| *e* | – | 1 | – | – | – | – | – |
| *ee* | – | 1 | – | – | – | – | – |
| *bb* | 2 | – | – | – | – | – | – |
| *m* | 1 | – | – | – | – | – | – |
| *dF* | – | – | – | 1 | – | – | – |
| *b”* | – | – | – | – | – | 1 | – |
| *e’* | 1 | 1 | – | – | 1 | – | – |
| *39* | – | 1 | – | – | – | – | – |
| *A* | 61 | 48 | 29 | 58 | 35 | 58 | 8 |
| *dD* | – | – | – | 1 | – | – | – |
| *b’* | – | – | 1 | – | – | – | – |