

DNA barcoding provides support for a cryptic species complex within the globally distributed and fishery important opah (*Lampris guttatus*)

JOHN R. HYDE,* KAREN E. UNDERKOFFLER† and MEAGAN A. SUNDBERG‡

*National Marine Fisheries Service, Southwest Fisheries Science Center, 8901 La Jolla Shores Dr., La Jolla, CA 92037, USA, †Joint Institute for Marine and Atmospheric Research, University of Hawaii, Honolulu, HI 96818, USA

Abstract

The cornerstone of fisheries management relies on a solid taxonomic base and an understanding of how animals can be grouped into coherent management units. Surprisingly, little is known about the basic biology and ecology of opah (*Lampris guttatus*), a globally distributed species that is commercially exploited and regionally common in the North Pacific. Recent efforts to collect life history data on this species uncovered evidence of two North Pacific morphotypes. Sequencing of the mitochondrial cytochrome c oxidase I gene (655 bp) for these morphotypes and other specimens collected worldwide ($n = 480$) produced five strongly diverged and well-supported clades. Additional sequence data from the cytochrome b gene (1141 bp) as well as the nuclear recombination activating gene 1 (1323 bp) corroborated these results, suggesting these five clades probably represent separate species. Our conclusion that opah is a complex of five separate species has implications for management and indicates a need to gather additional data on these poorly understood fishes.

Keywords: cryptic species, *Lampris guttatus*, moonfish, opah

Received 14 March 2014; revision received 15 April 2014; accepted 17 April 2014

Introduction

Though not a direct fishery target, the opah or moonfish (*Lampris guttatus*) is often captured in pelagic longline fisheries targeting bigeye tuna (*Thunnus obesus*) and broadbill swordfish (*Xiphias gladius*). While often classified as a nontarget species, the quality of its flesh and relative abundance in central North Pacific fisheries has made it a valuable component of commercial fisheries in this region with 196 047 individuals landed during the period 1994–2011 (Hawn & Collette 2012). The only other extant recognized species in the genus is *Lampris immaculatus*.

Despite its commercial value and common occurrence in pelagic fisheries, little is known regarding the life history of opah. Fossil data show that morphologically similar lamprid ancestors (*Megalampris keyesi*, Gottfried *et al.* 2006) have existed since at least the Oligocene (approximately 26 Ma) in southern latitudes and have been present off the coast of California since the late Miocene (approximately 6–10 Ma, *Lampris zatima*, Jordan & Gilbert 1920; Jordan 1920). To infer the life history and

habitat of this species, there have been a number of cursory studies including electronic tagging (Polovina *et al.* 2008), examination of unique eye-heater musculature (Runcie *et al.* 2009), diet studies (Polovina *et al.* 2008; Choy *et al.* 2013), early life history studies (Olney 1984), sexual dimorphism (Hawn *et al.* 2002), live colouration and maximum size (Hawn & Collette 2012), and preliminary age, growth and maturity studies (Francis *et al.* 2004). However, these studies were typically limited in sample size and geographic scope. In the course of collecting additional life history data on this species, observations were made by the authors and other researchers (Don (DH) Hawn, NMFS personal communication) that suggested the presence of two distinct morphotypes in the North Pacific. We subsequently initiated a study to investigate these differences and began collecting tissue samples for genetic analysis and compiling standard morphometric data.

Materials and methods

Sample collection

Fish were sampled opportunistically at the United Fishery Agency fish auction in Hawaii, by fishery observers

Correspondence: John R. Hyde, Fax: 858-546-7003; E-mail: John.Hyde@NOAA.gov

aboard commercial fishing vessels, from museum collections, and by collaborating fishers. In most cases, a sample of pectoral fin tissue was preserved in 95% undenatured ethanol; in some cases, white muscle or skin samples were obtained instead. The overwhelming majority of samples were obtained from the North Pacific (see Table 1); however, additional samples were obtained from throughout the species' known range to obtain a good representation of overall diversity (Fig. 1). Because a tissue specimen of the sister species *Lampris immaculatus* was unavailable, we used a sample of *Regalecus glesne* as an out-group and included *L. immaculatus* sequence data when available from GenBank.

DNA extraction and PCR

A chelex-based boiling protocol (Walsh *et al.* 1991; Hyde *et al.* 2005) was employed to extract DNA from all samples. The 'barcode' region of the mitochondrial cytochrome c oxidase I gene (COI) was amplified by polymerase chain reaction (PCR) using primers Fish_COI-F 5' TCW ACC AAC CAC AAA GAY ATY GGC AC and Fish_COI-R 5' TAR ACT TCW GGG TGR CCR AAG AAT CA (modified from Ward *et al.* 2005). A subset of samples had their mitochondrial cytochrome b (cyt b) gene and a portion of their nuclear recombinase activating gene 1 (rag1) amplified using primers Opah_t-Glu-F 5' TGGCTGAAAACACCCTTG with Opah_tThr-R 5' GGCGTCCAGCTTACAAGACT and Lampridae_Rag1-F 5' CCTGCCCTCAAGAAYGTGTC with Lampridae_Rag1-R 5'CGAAAACGGCGRAACAG TTTG, respectively. PCR was performed in 10 μ L volumes containing 67 mM Tris-HCl pH 8.8, 16.6 mM (NH₄)₂SO₄, 10 mM β -mercaptoethanol, 2 mM MgCl₂, 800 μ M dNTPs, 0.4 μ M each primer, 0.5 mg/mL bovine serum albumin, 0.5 units *Taq* DNA polymerase (New England Biolabs) and 50 ng of DNA template. The thermal profile for all reactions consisted of 92°C for 2 min, followed by 35 cycles of 92°C (20 s), 55°C (60 s) and 72°C

(60 s). No template, negative controls were run in all PCR sets to control for possible contamination of reagents. Products were electrophoresed through 2% agarose gel, stained with ethidium bromide and visualized on a UV transilluminator. PCR products were enzymatically cleaned using ExoSap-IT (Affymetrix). Cleaned products were sequenced in both directions using BIGDYE TERMINATOR v3.1 (Life Technologies) following manufacturer's protocols and analysed using an ABI3730 Genetic Analyzer (Life Technologies). Raw sequences were edited and aligned using SEQUENCHER v4.9 (GeneCodes).

Sequence analyses

Edited sequences were compiled and distance-based analyses were performed using MEGA v4.1 (Tamura *et al.* 2007). Pairwise distance measurements between individual sequences were performed using the Kimura 2-parameter (K2P) model. The K2P model was chosen for these analyses so as to be directly comparable to the Ward (2009) study which specifically addressed levels of divergence at the COI gene. The distances obtained were used to assess intraspecific and interspecific divergence and compared to published data generated for other taxa. Standard measures of genetic diversity, including nucleotide and gene diversity as well as counts of transitions and transversions, were performed using ARLEQUIN v3.1.1. (Excoffier *et al.* 2005).

Initial analyses of COI sequences produced several distinct and deeply diverged lineages. To further investigate this finding, additional sequence data were generated for the rag1 and cyt b genes from a subset of samples from each lineage. Evolutionary model testing for each gene was performed using JMODELTEST (Posada 2008). Phylogenetic relationships were evaluated for each gene using a Bayesian framework as implemented in BEAST v1.6.1 (Drummond & Rambaut 2007). Trees were rooted with *R. glesne* because this lampriform species was available for analysis. Whenever possible, sequence

Table 1 Regional sample sizes (*n*), number of unique haplotypes (*H*), haplotype diversity (*h*), nucleotide diversity (π), number of transitions (*T_S*), number of transversions (*T_V*), number of variable sites (*V*), uncorrected average pairwise differences between samples (*P*) and per cent nucleotide composition for cytochrome c oxidase I gene data

Location	<i>n</i>	<i>H</i>	<i>h</i> ± SD	π ± SD	<i>T_S</i>	<i>T_V</i>	<i>V</i>	<i>P</i> ± SD	% C	% T	% A	% G
North Pacific	408	66	0.894 ± 0.009	0.036 ± 0.018	115	19	125	23.91 ± 10.53	31.5	27.0	22.2	19.3
Indian	12	7	0.773 ± 0.128	0.039 ± 0.021	44	7	49	25.52 ± 12.06	31.4	27.1	22.3	19.3
South Pacific	21	8	0.719 ± 0.099	0.002 ± 0.001	6	1	7	1.01 ± 0.71	31.8	26.7	22.4	19.1
South Atlantic	25	15	0.907 ± 0.045	0.019 ± 0.010	48	8	52	12.33 ± 5.76	31.7	26.8	22.4	19.2
North Atlantic	4	4	1.000 ± 0.177	0.063 ± 0.042	60	11	71	41.33 ± 22.94	31.4	27.2	22.6	18.9
S. Atlantic/Indian*	9	6	0.889 ± 0.091	0.003 ± 0.002	6	1	7	2.22 ± 1.35	31.8	26.7	22.4	19.1
Mediterranean	1	1	N/A	N/A	N/A	N/A	N/A	N/A	31.6	27.0	22.9	18.5

*These samples lacked locality data other than they came from either of these two regions.

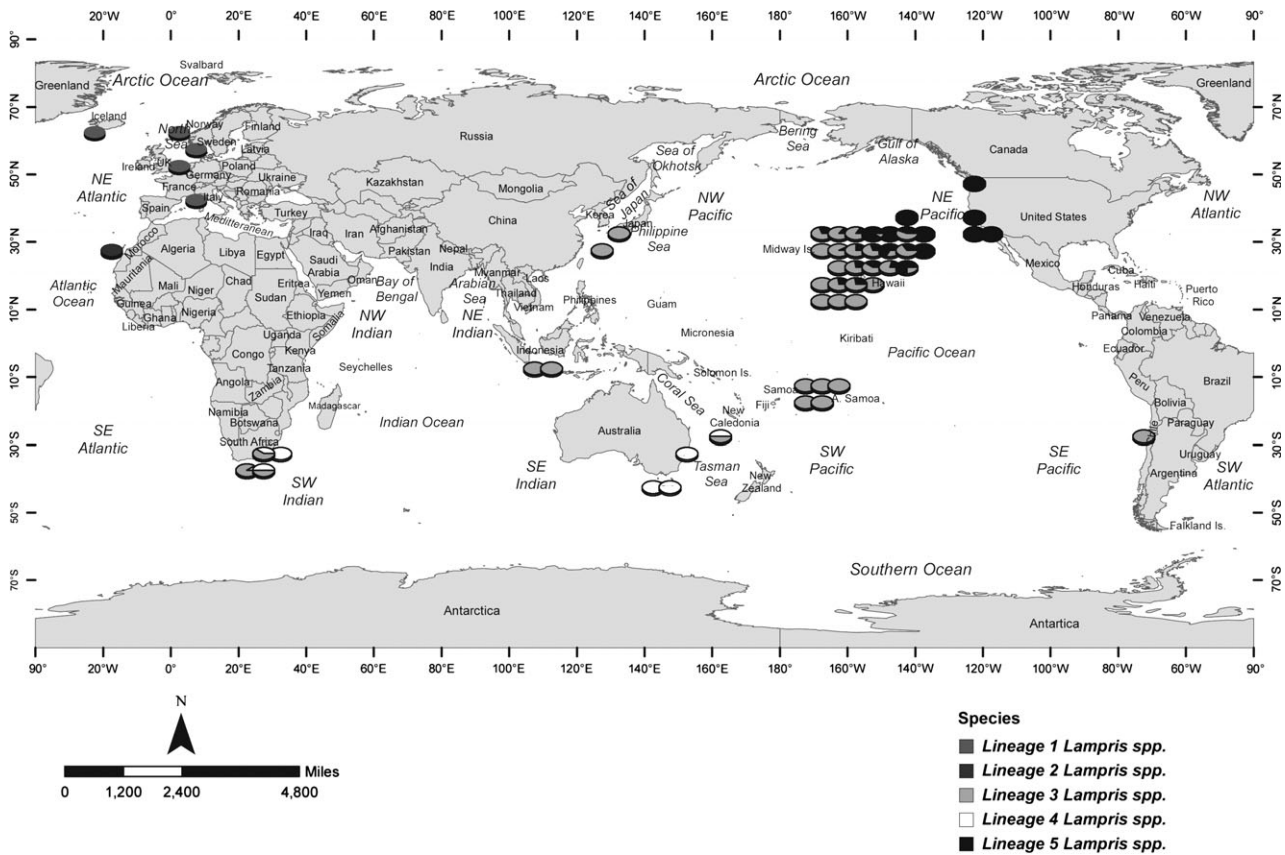


Fig. 1 Distribution of samples analysed in this study in 10° blocks. Pie charts are coded to represent the composition of cytochrome oxidase I gene lineages present in each block.

data from the sister taxon *L. immaculatus* were included in the analyses.

Results

Analysis of 655 base pairs (bp) of sequence data for the COI gene from 480 *Lampris guttatus* specimens produced 82 unique haplotypes (GenBank accession# JF931865–JF931947). Overall genetic diversity was exceptionally high with 140 polymorphic sites and overall nucleotide diversity of $\pi = 0.038$ (Table 2). These high levels of diversity were largely retained when samples were broken into oceanographic regions (Table 1). Per cent

divergence (K2P) between individual *L. guttatus* sequences ranged from 0% to 13.2%, divergence between *L. guttatus* and a single *L. immaculatus* sequence ranged from 12.0% to 14.4%. Levels of divergence were not normally distributed and instead presented as several discrete modes (Fig. 2A). Even though only a small subset of these samples ($n = 19$) were analysed at the other genes, the high diversity was largely retained (Table 2) with per cent divergence for *cyt b* (13 unique haplotypes, GenBank accession# JF931948–JF931967), and *rag1* (14 unique genotypes, GenBank accession# JF931968–JF931987), sequences ranging from 0% to 21.4% (Fig. 2B) and 0% to 0.69% (Fig. 2C), respectively. This variation in

Table 2 Molecular diversity measures for analysed genes with overall sample sizes (n), number of unique haplotypes (H), haplotype diversity (h), nucleotide diversity (π), number of transitions (T_s), number of transversions (T_v), number of variable sites (V), uncorrected average pairwise differences between samples (P), per cent nucleotide composition and total base pairs analysed (bp)

Gene	n	H	$h \pm SD$	$\pi \pm SD$	T_s	T_v	V	$P \pm SD$	% C	% T	% A	% G	bp
COI	480	82	0.896 ± 0.009	0.038 ± 0.018	131	22	140	24.65 ± 10.84	31.6	27.0	22.2	19.3	655
<i>cyt b</i>	19	13	0.947 ± 0.033	0.090 ± 0.045	245	80	286	102.39 ± 46.04	33.3	27.8	23.0	16.0	1141
<i>rag1</i>	19	14	0.964 ± 0.028	0.005 ± 0.003	6	7	13	7.12 ± 3.50	23.9	23.1	25.5	27.5	1323

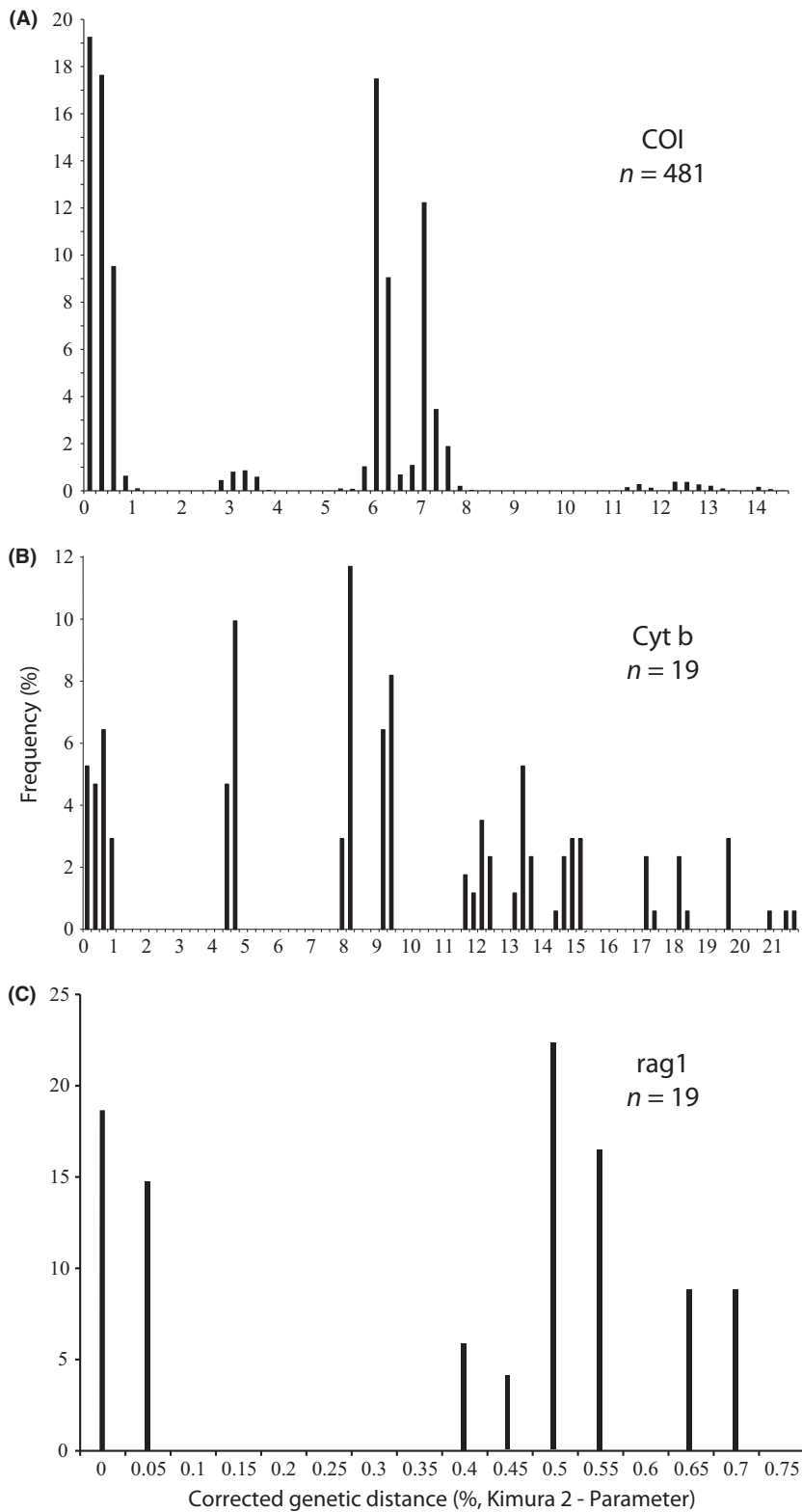


Fig. 2 Frequency plot of pairwise distances (K2P) between cytochrome c oxidase I gene (COI) (A), cyt b (B) and rag1 (C) sequences from individual *Lampris guttatus* samples used in this study. COI data include a single sequence from *L. immaculatus* (GenBank DQ108066), pairwise distance from *L. guttatus* sequences ranged from 12.0% to 14.4%.

evolutionary rate among these genes is in line with other studies that have shown COI evolves at approximately 60–70% the rate of the cyt b gene but much faster than

the rag genes (e.g. Hyde & Vetter 2007). As is typical for genes that code for functional proteins, the majority of nucleotide substitutions were at 3rd codon positions

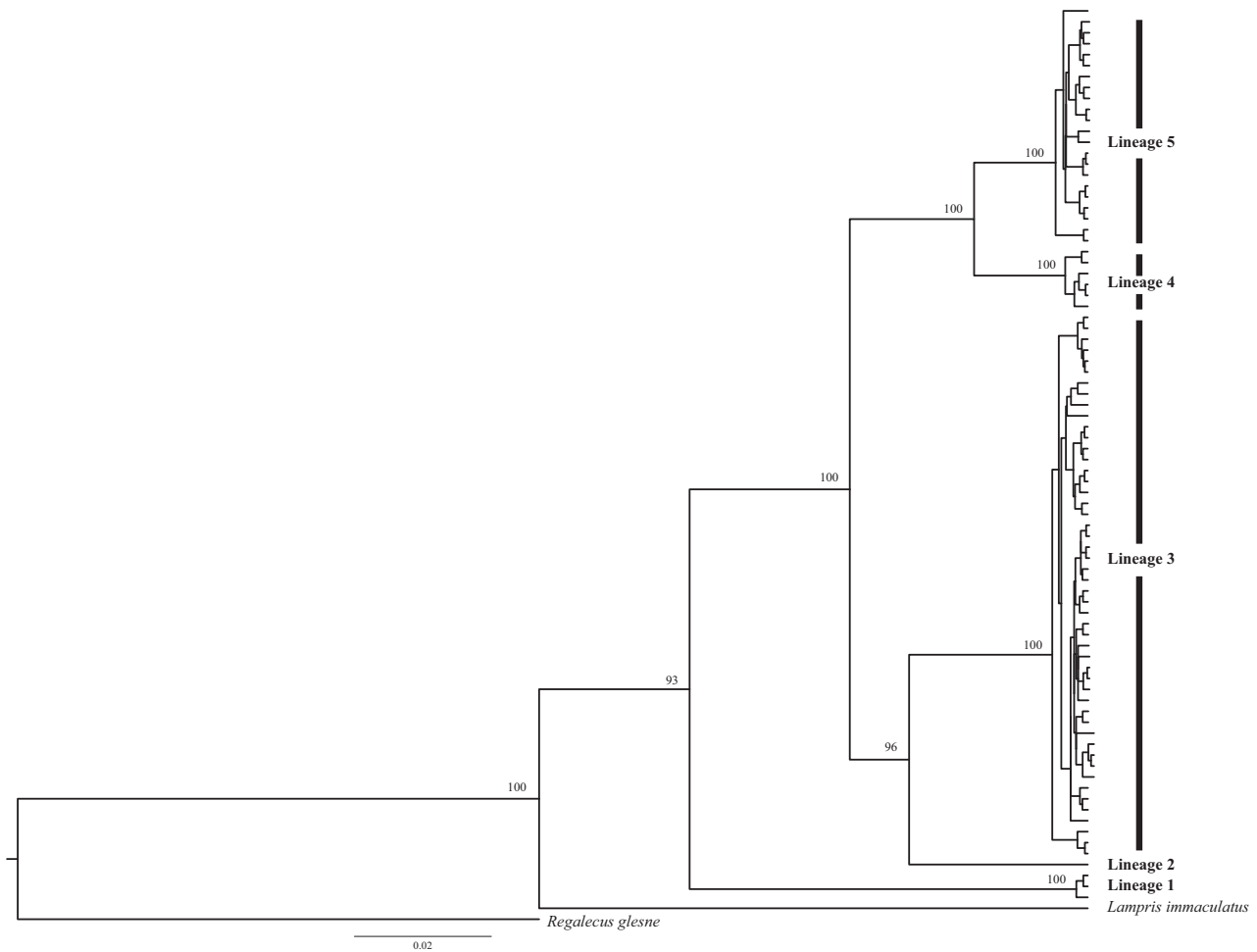


Fig. 3 Bayesian-derived tree with maximum posterior probability for cytochrome c oxidase I gene haplotypes found in this study. Values above nodes represent posterior support values. Lineage names for each of the clades present in this figure are referred to in Figs 4 and 5. Sequence data from *Lampris immaculatus* (GenBank DQ108066) and *Regalecus glesne* (SIO 97–226) are used as out-groups.

with no resultant change in amino acid sequence. The majority of substitutions that resulted in amino acid changes were confined to interlineage comparisons.

Evolutionary model testing indicated the general time-reversible (GTR) model considering a proportion of invariant sites (I) and gamma distribution (G)

(Rodriguez *et al.* 1990) had the highest likelihood for each of the three data sets (COI, cyt b, rag1) and was used for construction of Bayesian-derived trees using BEAST. The tree with maximum posterior probability for COI data produced five deeply diverged and strongly supported clades within *L. guttatus* (Fig. 3). Individuals

Table 3 Average per cent divergence (Kimura 2–P) between cytochrome c oxidase I gene (COI) haplotype clades below diagonal, net divergence above diagonal, average within clade haplotype divergence in bold. Lineages containing only a single specimen are denoted NA for within lineage distance

	Lineage 1	Lineage 2	Lineage 3	Lineage 4	Lineage 5	<i>Lampris immaculatus</i>	within lineage
Lineage 1	*	11.55	12.35	10.73	11.17	13.87	0.31
Lineage 2	11.70	*	5.57	6.12	5.71	11.85	NA
Lineage 3	12.71	5.78	*	6.95	6.65	13.89	0.41
Lineage 4	11.14	6.37	7.41	*	2.68	12.73	0.50
Lineage 5	11.55	5.93	7.08	3.15	*	12.17	0.45
<i>Lampris immaculatus</i>	14.02	11.85	14.10	12.98	12.39	*	NA

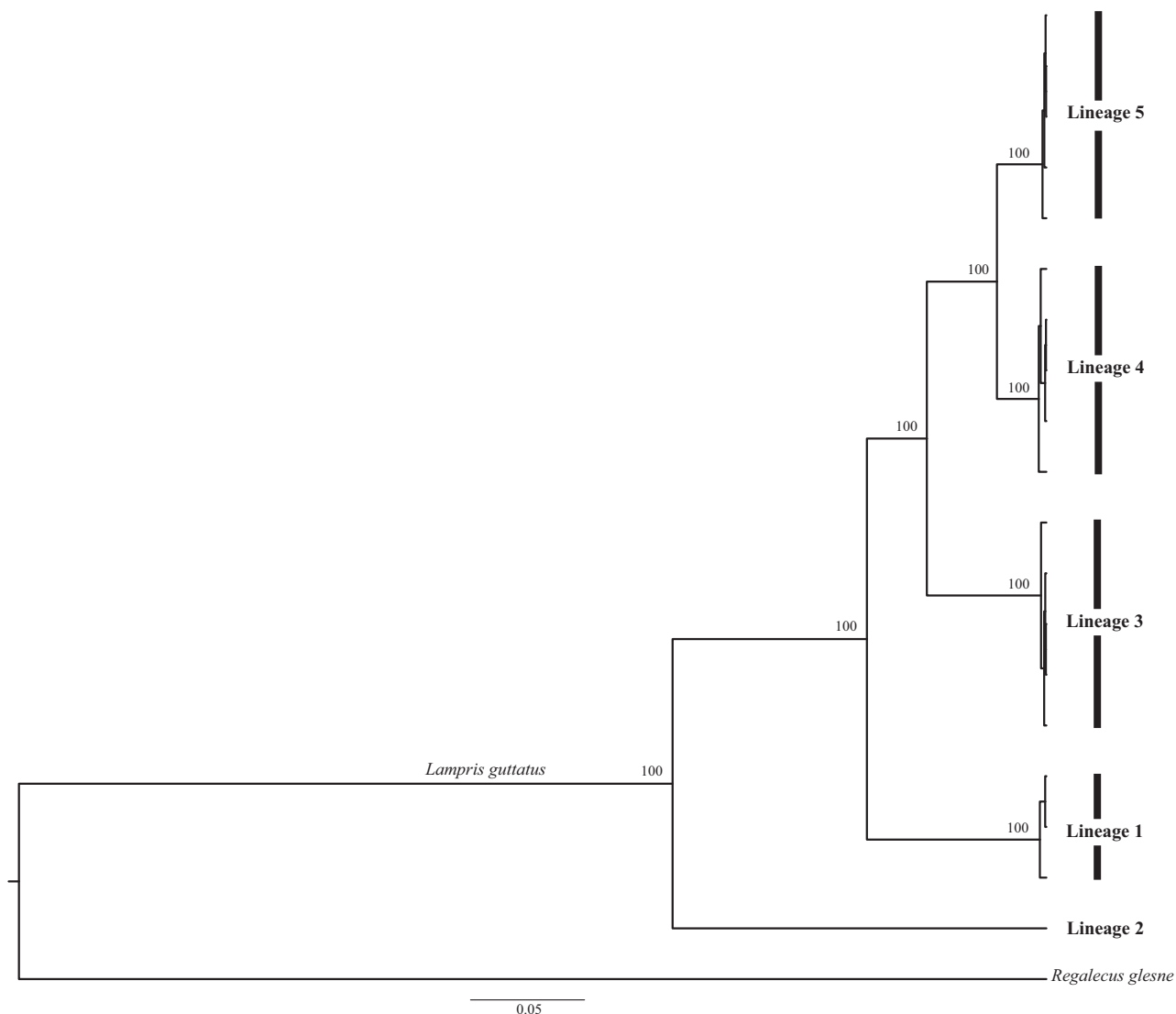


Fig. 4 Bayesian-derived tree with maximum posterior probability for cyt b sequences generated in this study. Values above nodes represent posterior support values. Sequences are labelled with cytochrome c oxidase I gene lineage from Fig. 3. A single sequence from *Regalecus glesne* (SIO 97–226) is used as an out-group.

within each clade were subdivided and further analysed by these unique COI lineages. Interclade divergence (K2P) ranged from 2.7% to 12.4% while values for intraclade divergence ranged from 0% to 0.9% with average intraclade haplotype divergence ranging from 0.3% to 0.5% (Table 3).

Sequence data from *rag1* and *cyt b* for individuals from each COI lineage produced highly similar trees to that derived from the COI data (Figs 4 and 5). In all cases, there was a 100% concordance between COI clades and those generated for the other two genes; high posterior support resulted in all cases. The placement of Lineage 2 relative to the other lineages was the only structural difference noted between gene trees. Such discrepancies between gene trees are common and can be

caused by multiple factors including incomplete lineage sorting, selective or constraining forces acting upon individual genes and hybridization (e.g. Maddison 1997).

Discussion

Ward (2009) found that 92% of conspecific fishes showed <1% haplotype divergence at the barcode region of the COI gene, while 98% showed divergence <2%. Based on these findings, he suggested that intraspecific divergence >2% probably characterized cryptic species complexes. Pairwise distances between COI haplotypes for *L. guttatus* ranged from 0% to 13.2% with the presence of five distinct clades with interclade divergences of 2.7% to 12.4%. Values for intraclade divergence ranged from 0%

to 0.9%. Following Ward (2009), the data suggest that these individual clades represent cryptic species as all interclade divergences are >2% and intraclade divergences are <1%.

To further evaluate the COI findings, we sequenced a subset of samples representing members of the different COI clades, for an additional mitochondrial gene (cyt b) and a nuclear gene (rag1). Trees for all three genes produced the same five distinct clades (Figs 3–5) with strong posterior support. Concordance of the two mitochondrial genes and one nuclear gene strongly support that the five clades found in the COI analysis are on separate evolutionary trajectories; given current divergence and comparison to other fishes (Ward 2009), all five warrant separate species status based upon the genetic data and forthcoming evidence of distinguishing morphological characters (KEU & MAS unpublished data).

Genetic evaluation of *L. guttatus* was initiated due to the presence of two morphotypes observed in the North

Pacific. These morphotypes corresponded 100% with two of the COI clades found in this study (Lineages 3 and 5) and are approximately 6.7% diverged at this gene. The morphotypes were originally categorized by relative eye size, but subsequent detailed morphologic study has uncovered several other diagnostic characters that differentiate these fishes (KEU & MAS unpublished data). Although fewer specimens from the other lineages have been available for study, initial data support the conclusion that significant morphologic distinctions exist among them as well. The agreement between morphology and genetics, especially because several of these lineages occur in sympatry (see Fig. 1), adds further support to the assertion that these are separate species. The genetic and morphological data generated in this study are in the process of being combined into formal descriptions for several of the species.

The finding of a cryptic species complex in a widely distributed marine fish that is well-known or

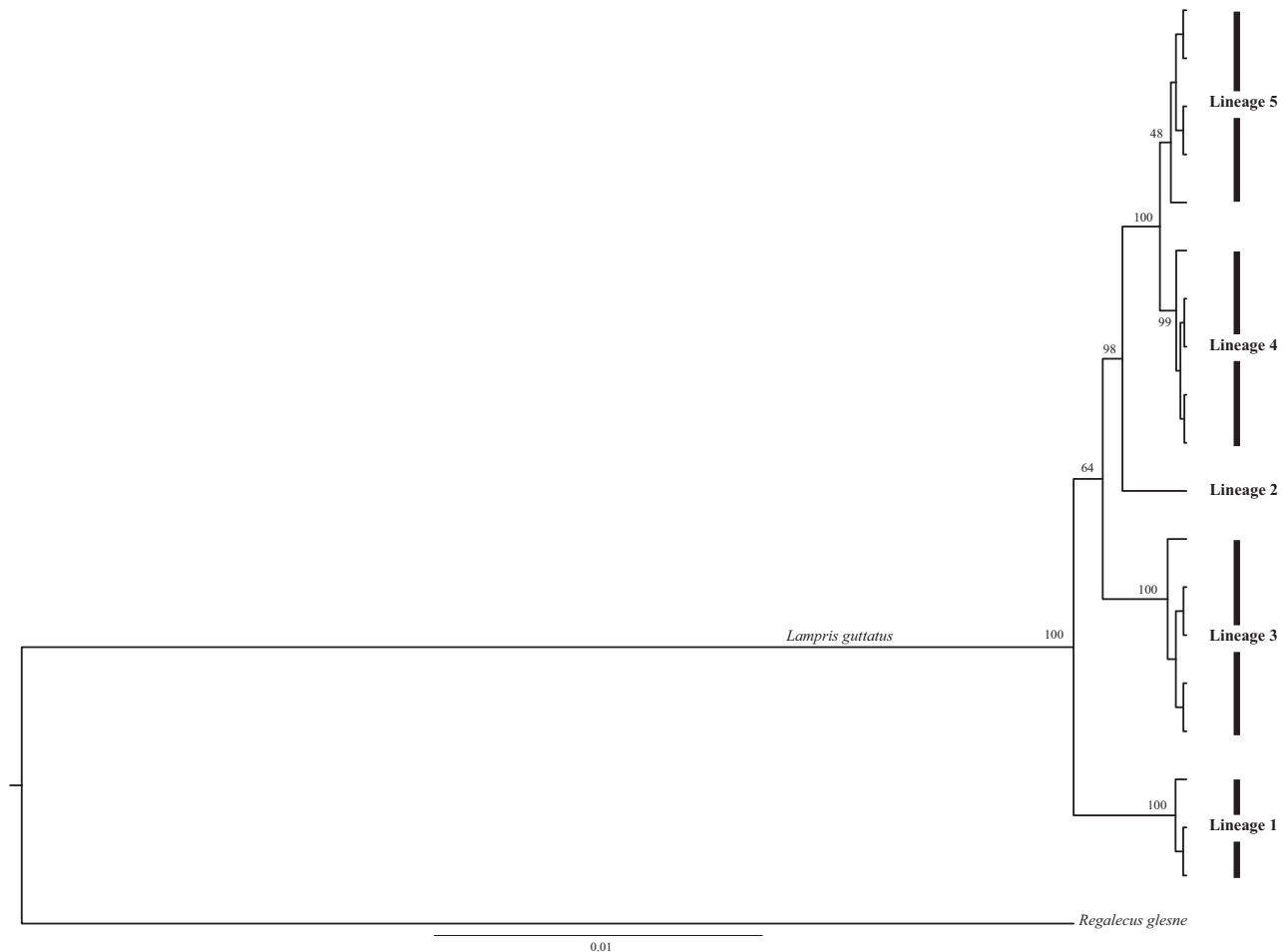


Fig. 5 Bayesian-derived tree with maximum posterior probability for rag1 sequences generated in this study. Values above nodes represent posterior support values. Sequences are labelled with cytochrome c oxidase I gene lineage from Fig. 3. A single sequence from *Regalecus glesne* (SIO 97–226) is used as an out-group.

an important fishery component is not unique (e.g. Brendtro *et al.* 2008; Hyde *et al.* 2008; Yoshita *et al.* 2009) and may be more common in poorly studied pelagic mid-water species (e.g. Craig *et al.* 2004; Roberts 2012). Review of literature and sequences deposited in BOLD and GenBank suggest that several lampridiform genera contain cryptic lineages and are probably in need of taxonomic revision. Following detailed genetic and morphologic comparisons, the crestfish genus *Lophotus* was found to contain at least two species in the Pacific with uncertainty as to the status of fish in other ocean basins (Craig *et al.* 2004). A thorough review by Roberts (2012) offers multiple lines of evidence for existence of at least two species within *Regalecus*. Comparisons of conspecific sequences for *Zu cristatus* and *Metavelifer multiradiatus* available from the GenBank and BOLD databases show evidence for strongly diverged lineages.

Although mining from online databases has been shown to yield false levels of diversity due to species misidentifications and sequence errors (Harris 2003), these hypotheses can be tested if voucher specimens archived in permanent museum collections can be examined. Much of the taxonomic confusion within this group of fishes is likely due to the lack of type specimens available and relative rarity of capture, which has led to a preponderance of synonyms for several species including opah. In the case of opah, although at least six synonyms currently exist for *L. guttatus* (*L. gunneri*, *L. imperialis*, *L. lauta*, *L. luna*, *L. regius*, *L. stroemii*), a complete lack of type specimens makes reappraisals of synonymy difficult. Application of genetic barcoding techniques to many of these wide-ranging species will probably yield additional examples of cryptic diversity, better clarify the taxonomy of this group of fishes and add to our understanding of speciation in the mesopelagic realm.

Implications for management

Even though no lampridiform species is at present directly targeted by commercial fisheries, the opah has become an important component in some fisheries, particularly those targeting bigeye tuna and swordfish in the North Pacific. The value of US commercial landings of opah has risen rapidly in recent years from <\$4000 in 1999 to almost \$2.5 million in 2009. Results from this study provide clear evidence that *L. guttatus* is not a single, globally distributed species. The overwhelming majority (99.9%) of US landings of opah occur in Hawaii where it seems at least two species (Lineages 3 and 5) interact with fisheries, though one dominates the catch in this region (Lineage 3, approximately 90%, KEU & MAS unpublished data). However, initial monitoring of

catch in California waters shows that the species composition is strongly skewed towards the other species (Lineage 5, >95%, JRH unpublished data). Clearly, further work needs to be done to delineate the distribution, habitat and life history of these cryptic lineages and how they interact with world fisheries to improve future management plans for this growing fishery. Further, previous studies of life history, diet and habitat will need to be re-evaluated in the light of these findings.

Acknowledgements

Tissue samples were graciously provided by B. Takenaka, N. Arberilla (UFA, Hawaii); B. Luduc (Euronor, France); J. Pals-son (Marine Research Institute, Iceland); S. Kullander, B. Delling, E. Ahlander (Swedish Museum of Natural History, Sweden); M. Bruni (Oceanographic Museum of Monaco, France); A. Hay, S. Reader (Australian Museum, Australia); J. Friel (Cornell University, USA); D. Nelson, A. Belasen (University of Michigan Museum of Zoology, USA); A. Connell, R. Bills (South Africa Institute for Aquatic Biodiversity, South Africa); S. Lamberth (Marine and Coastal Management, South Africa); A. Graham (CSIRO, Australia); S. Kostelnik and participating American Samoa longline observers (NMFS, USA); D. Roje (University of Washington, USA); A. de Vera Hernandez (Museo de Ciencias Naturales de Tenerife, Spain); K. Busscher, J. Kelly, M. Marsik, and all participating Hawaii Longline observers and debriefers (NMFS, USA); G. Shinahara (National Museum of Nature and Science, Japan); J. Clayton, J. Finan, D. Johnson, and J. Williams (National Museum of Natural History, USA); H. J. Walker (Scripps Institution of Oceanography, USA). We are grateful to E. DeMartini, A. Lang, and B. Mundy for providing comments on earlier drafts of this manuscript and G. Kitazaki (Diamond Head Seafood, Hawaii) for pointing out morphological variations in opah specimens to KEA and MAS during UFA visits.

References

- Brendtro KS, McDowell JR, Graves JE (2008) Population genetic structure of escolar (*Lepidocybium flavobrunneum*). *Marine Biology*, **155**, 11–22.
- Choy CA, Portner E, Iwane M, Drazen JC (2013) Diets of five important predatory mesopelagic fishes of the central North Pacific. *Marine Ecology Progress Series*, **492**, 169–184.
- Craig MT, Hastings PA, Pondella DJ (2004) Notes on the systematics of the crestfish genus *Lophotus* (Lampridiformes: Lophotidae), with a new record from California. *Bulletin of the Southern California Academy of Sciences*, **103**, 57–65.
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Francis MP, Griggs L, Maolagáin CÓ (2004) *Growth Rate, Age at Maturity, Longevity and Natural Mortality Rate of Moonfish (Lampris guttatus)*. Final Research Report for Ministry of Fisheries Research Project TUN2003-01 Objective 1. National Institute of Water and Atmospheric Research, New Zealand. Available from <http://fs.fish.govt.nz/Page.aspx?pk=113&dk=22484> (accessed 27 February 2012).

- Gottfried MD, Fordyce RE, Rust S (2006) *Megalompris keyesi*, a giant moonfish (Teleostei, Lampridiformes) from the late Oligocene of New Zealand. *Journal of Vertebrate Paleontology*, **26**, 544–551.
- Harris DJ (2003) Can you bank on GenBank? *Trends in Ecology and Evolution*, **18**, 317–319.
- Hawn DR, Collette BB (2012) What are the maximum size and live body coloration of opah (Teleostei: Lampridae: Lampris species)? *Ichthyological Research*, **59**, 272–275.
- Hawn DR, Seki M, Nishimoto R (2002) *An investigation of the life history and ecology of opah and monchong in the North Pacific*. Report submitted to the 15th Meeting of the Standing Committee on Tuna and Billfish. Honolulu, Hawaii. July 2002. Available from <http://www.soest.hawaii.edu/PFRP/sctb15/papers/BBRG-2.pdf> (accessed 27 February 2012).
- Hyde JR, Vetter RD (2007) The origin, evolution, and diversification of rockfishes of the genus *Sebastes* (Cuvier). *Molecular Phylogenetics and Evolution*, **44**, 490–811.
- Hyde JR, Lynn E, Humphreys R Jr, Musyl M, West AP, Vetter R (2005) Shipboard identification of fish eggs and larvae by multiplex PCR, and description of eggs of blue marlin, shortbill spearfish, and wahoo. *Marine Ecology Progress Series*, **286**, 269–277.
- Hyde JR, Kimbrell CA, Budrick JE, Lynn EA, Vetter RD (2008) Cryptic speciation in the vermilion rockfish (*Sebastes miniatus*) and the role of bathymetry in the speciation process. *Molecular Ecology*, **17**, 1122–1136.
- Jordan DS (1920) An ancient moonfish. *The Scientific Monthly*, **11**, 470–473.
- Jordan DS, Gilbert JZ (1920) *Fossil Fishes of the Diatom Beds of Lompoc*, 45 pp. Stanford University Publication, Stanford, California.
- Maddison WP (1997) Gene trees in species trees. *Systematic Biology*, **46**, 523–536.
- Olney JE (1984) Lampridiformes: development and relationships. In: *Ontogeny and Systematics of Fishes* (eds Moser HG, Richards WJ, Cohen DM, Fahay MP, Kendall AW Jr & Richards SL), pp. 368–379. American Society of Ichthyologists and Herpetologists Special Publication No. 1, Lawrence, Kansas.
- Polovina JJ, Hawn D, Abecassis M (2008) Vertical movement and habitat of opah (*Lampris guttatus*) in the central North Pacific recorded with pop-up archival tags. *Marine Biology*, **153**, 257–267.
- Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.
- Roberts TR (2012) *Systematics, Biology, and Distribution of the Species of the Oceanic Oarfish Genus Regalecus (Teleostei, Lampridiformes, Regalecidae)*, 268 pp. Mémoires du Muséum National d'Histoire Naturelle, Paris.
- Rodriguez R, Oliver JL, Marin A, Medina JR (1990) The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology*, **142**, 485–501.
- Runcie RM, Dewar H, Hawn DR, Frank LR, Dickson KA (2009) Evidence for cranial endothermy in the opah (*Lampris guttatus*). *Journal of Experimental Biology*, **212**, 461–470.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.
- Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques*, **10**, 506–513.
- Ward RD (2009) DNA barcode divergence among species and genera of birds and fishes. *Molecular Ecology Resources*, **9**, 1077–1085.
- Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B*, **360**, 1847–1857.
- Yoshita Y, Yamanoue Y, Sagara K *et al.* (2009) Phylogenetic relationship of two *Mola* sunfishes (Tetraodontiformes: Molidae) occurring around the coast of Japan, with notes on their geographical distribution and morphological characteristics. *Ichthyological Research*, **56**, 232–244.

J.R.H., K.E.U., and M.A.S. designed the study. K.E.U. and M.A.S. collected specimens and morphometric data. J.R.H. obtained, edited and analysed genetic data. All authors contributed to the final version of this manuscript.

Data Accessibility

DNA sequences: GenBank accession numbers JF931865–JF931987. Sequence alignments, tree files, sample information: Dryad entry doi:10.5061/dryad.71m33.