

Hidden Diversity in Sardines: Genetic and Morphological Evidence for Cryptic Species in the Goldstripe Sardinella, *Sardinella gibbosa* (Bleeker, 1849)

Rey C. Thomas Jr.^{1*}, Demian A. Willette², Kent E. Carpenter³, Mudjekeewis D. Santos¹

1 Genetic Fingerprinting Laboratory, National Fisheries Research and Development Institute, Quezon City, Metro Manila, Philippines, **2** Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, California, United States of America, **3** Department of Biological Sciences, Old Dominion University, Norfolk, Virginia, United States of America

Abstract

Cryptic species continue to be uncovered in many fish taxa, posing challenges for fisheries conservation and management. In *Sardinella gibbosa*, previous investigations revealed subtle intra-species variations, resulting in numerous synonyms and a controversial taxonomy for this sardine. Here, we tested for cryptic diversity within *S. gibbosa* using genetic data from two mitochondrial and one nuclear gene regions of 248 individuals of *S. gibbosa*, collected from eight locations across the Philippine archipelago. Deep genetic divergence and subsequent clustering was consistent across both mitochondrial and nuclear markers. Clade distribution is geographically limited: Clade 1 is widely distributed in the central Philippines, while Clade 2 is limited to the northernmost sampling site. In addition, morphometric analyses revealed a unique head shape that characterized each genetic clade. Hence, both genetic and morphological evidence strongly suggests a hidden diversity within this common and commercially-important sardine.

Citation: Thomas RC Jr, Willette DA, Carpenter KE, Santos MD (2014) Hidden Diversity in Sardines: Genetic and Morphological Evidence for Cryptic Species in the Goldstripe Sardinella, *Sardinella gibbosa* (Bleeker, 1849). PLoS ONE 9(1): e84719. doi:10.1371/journal.pone.0084719

Editor: Erik Sotka, College of Charleston, United States of America

Received: August 5, 2013; **Accepted:** November 18, 2013; **Published:** January 8, 2014

Copyright: © 2014 Thomas et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This project was funded mainly through a National Fisheries Research and Development Institute grant to Mudjekeewis Santos, Demian Willette and Eunice Bognot. This research was partly supported by a grant from the National Science Foundation (NSF OISE-0730256) to Kent Carpenter and Paul Barber. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: reychthomas@gmail.com

Introduction

Accurately defining the limits of a species is essential to studying its biology, ecology, conservation, and management [1–3]. Species are typically categorized according to gross morphology [4]. However, species diversity can be masked by a lack of obvious morphological differences between cryptic species [5]. Cryptic species are morphologically similar but genetically distinct lineages, and are often overlooked by species identification using gross morphology alone [6]. Genetic differentiation can be used to distinguish morphologically similar lineages [7–9], in which the genotypic-clustering species definition is utilized to identify cryptic species [10–12]. Cryptic species are widely distributed across different taxa and geographies in the marine realm [13–14], occurring either in allopatry [15,16] or sympatry [8,17] as sister species [16], or a result of convergent evolution [18]. In the Indo-West Pacific, cryptic species of widely distributed reef fishes contribute significantly to overall marine biodiversity [19]. Thus, inaccurate species delimitation would overlook cryptic species and underestimate biodiversity, a fact that can lead to flawed conservation and management strategies [5,20].

Marine small pelagic fishes comprise the majority of world's landed fish catch [21], and members of the family Clupeidae contribute significantly to this volume [22]. The Clupeidae reach their highest diversity in the Indo-West Pacific [23], a region also

proposed to be the geographic origin of this family [24]. Within the Indo-West Pacific, the goldstripe sardinella *Sardinella gibbosa* (Bleeker, 1849) is among the most abundant and widespread marine pelagic species. It has a distribution that extends from the East African coast to Taiwan, the Philippines, Indonesia and Northern Australia [23]. Historically, it is among the most abundant and commercially important species in the Indo-West Pacific sardine fishery [25,26]. In particular, the goldstripe sardinella is the second most abundant sardine occurring in the Philippine archipelago [27]. Migratory patterns of *S. gibbosa* have been correlated with the availability and seasonality of planktonic prey in the environment [26,28,29]. Biological data from the coast of India suggests a peak spawning period that lasts from early March towards the end of May [26,30]. However, two distinct length and age groups have been observed [31], and variations in the number of scale striations have been observed in *S. gibbosa* off the South African coast [32]. Morphological classification of this sardine has been complicated by subtle intra-species variations, leading to several recorded synonyms for *S. gibbosa* [23]. Type specimens, now considered synonyms of *S. gibbosa*, were previously known as *Clupea immaculata* (Southern Japan and China) [33], *Fimbriclupea dactyolepis* (Northwest Australia) [34], and *Sardinella taiwanensis* (Taiwan) [35]. Such subtle biological and morphological differences documented in *S. gibbosa* may hint of hidden diversity within the sardine.

Table 1. Sampling information for *S. gibbosa*.

Location	N	Collection date
Sta. Ana, Cagayan (CAG)	50	June, 2012
Atimonan, Quezon (QUE)	30	July, 2011
Balayan Bay, Batangas (BAT)	22	November, 2011
Manila Bay, Manila (MNL)	30	May, 2012
Tacloban, Leyte (LEY)	30	July, 2011
Ilo-Ilo City, Ilo-Ilo (ILO)	30	May, 2011
Banate, Ilo-Ilo (BAN)	32	April, 2013
Puerto Princesa, Palawan (PAL)	24	December, 2012
Kudat, Sabah (KUD)	21	March, 2011
Yilan County, Taiwan (YIL)	1	April, 2011
Nha Trang, Vietnam (NTR)	26	October, 2011
Phu Quoc, Vietnam (PQU)	26	October, 2011
Songkla, Thailand (SON)	4	October, 2011
Surat Thani, Thailand (SUR)	17	October, 2011
Koh Samui, Thailand (KOH)	8	October, 2011
Trang Province, Thailand (TRA)	27	October, 2011
Total	378	

doi:10.1371/journal.pone.0084719.t001

The objective of this work has been to explore the possible cryptic species within the goldstripe sardinella in the Philippine archipelago by examining molecular and morphometric data. We investigated the occurrence of cryptic species using genotypic clustering for both mitochondrial and nuclear markers. We also

used morphometric variations to test for subtle morphological differentiation between the genetically-partitioned groups. Both genetic and morphological evidence provide strong support for an unexpected multiple-species complex within the common and commercially-important sardine *S. gibbosa* in the Philippine archipelago.

Materials and Methods

Sampling

A total of 378 individuals of *S. gibbosa* were collected from sixteen fish markets across the Philippines, Taiwan, Malaysia, Vietnam and Thailand (Table 1). Body coloration, morphometric, and meristic characters were recorded for frozen then thawed samples. Tissue samples and voucher specimens preserved in absolute ethanol were stored at the National Fisheries Research and Development Institute, Quezon City, Philippines.

PCR amplification

Genomic DNA was extracted from muscle tissue samples using either modified Chelex® (Bio-Rad, Hercules, CA) DNA extraction protocol [36], or salting-out method [37]. Approximately 540 bp of the ribosomal 16S gene region are amplified by polymerase chain reaction (PCR) using the primers 16Sar (5'-CGCCTGTTTATCAAAAACAT-3') and 16Sbr (5'-CCGGTCTGAACTCAGATCACGT-3') [38]. Additional mitochondrial and nuclear DNA sequences were obtained only for Philippine collections due to mounting sequencing costs and limited time available for this study. The primers CRA (5'-TTCCACCTCTAACTCCCAAAGCTAG-3') and CRE (5'-CCTGAAGTAGGAACCAGATG-3') were used to amplify 560 bp of mitochondrial DNA control region [39].

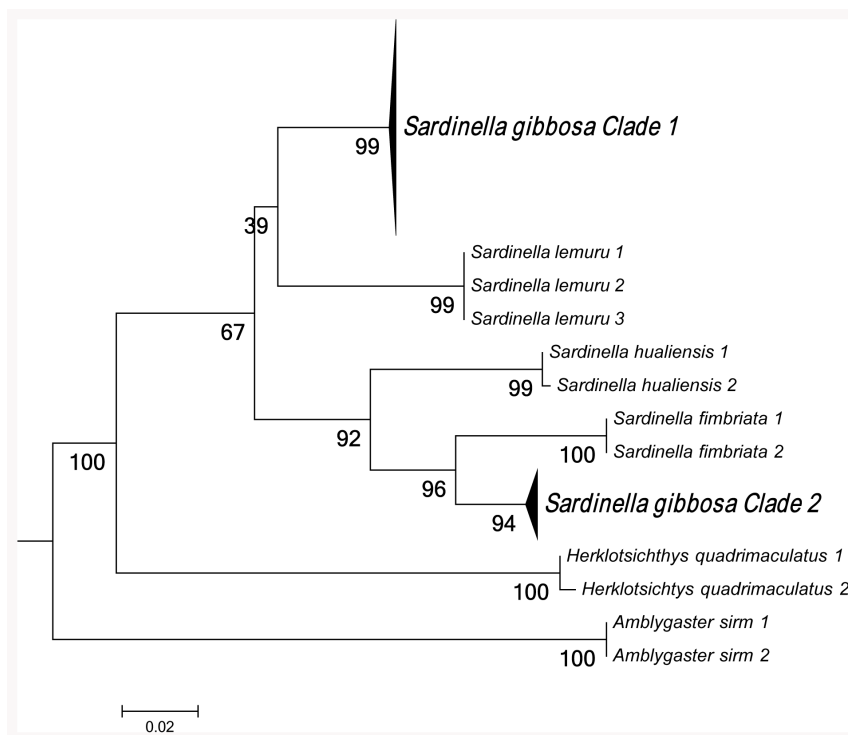


Figure 1. Maximum-likelihood (ML) tree inferred using 16S rRNA sequences. Bootstrap support values were calculated using 1,000 replicates. Sequences of *S. hualiensis* obtained from GenBank were JN580490.1 and JN580479.1.
doi:10.1371/journal.pone.0084719.g001

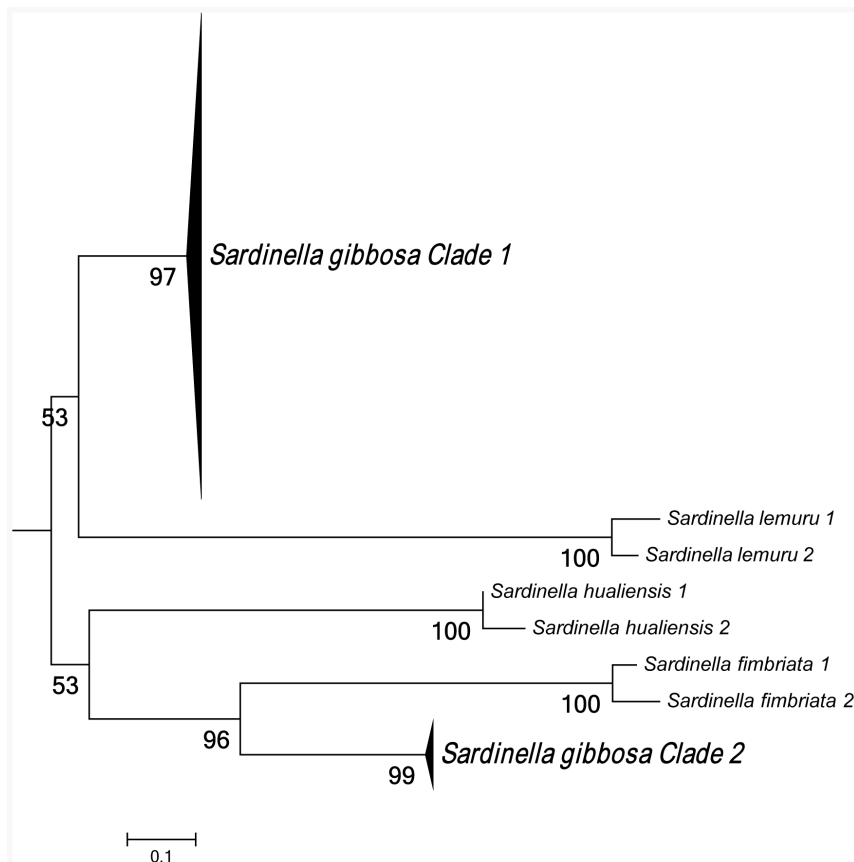


Figure 2. Maximum-likelihood (ML) tree generated from mitochondrial control region sequences. Bootstrap support values were calculated using 1,000 replicates.
doi:10.1371/journal.pone.0084719.g002

Nuclear DNA was obtained using the primers S7RPEX1F (5'-TGGCCTCTTCCTTGGCCGTC-3') and S7RPEX2R (5'-AACTCGTCTGGCTTTTCGCC-3') to amplify 700 bp of the 1st intron of the ribosomal S7 gene [40]. The total volume of

the reaction mixture to amplify mtDNA gene regions is 25 μ l; consisting of 13.5 μ l of nuclease-free water, 2.5 μ l of 10x PCR buffer, 2.5 μ l of 10 mM dNTP, 2.0 μ l of 25 mM MgCl₂, 1.0 μ l of 5x BSA, 1.25 μ l of 10 mM of both primers, and 0.125 μ l of *Taq*

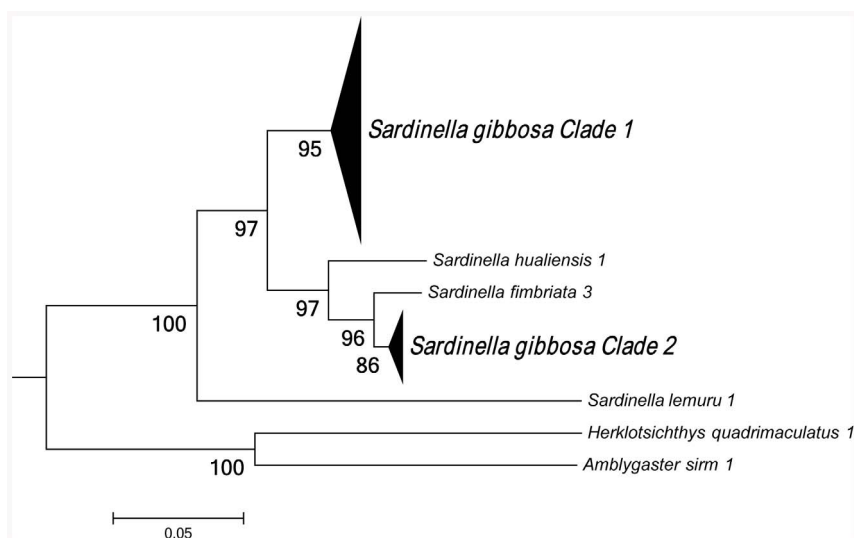


Figure 3. Maximum-likelihood (ML) tree generated from the 1st intron of ribosomal S7 sequences. Bootstrap support values were calculated using 1,000 replicates.
doi:10.1371/journal.pone.0084719.g003

Table 2. Pairwise genetic distances calculated for each taxa.

16S							
Species	1	2	3	4	5	6	7
<i>Sardinella gibbosa</i> Clade 1		0.03	0.05	0.04	0.03	0.08	0.07
<i>Sardinella gibbosa</i> Clade 2	0.13		0.02	0.03	0.04	0.09	0.10
<i>Sardinella fimbriata</i>	0.22	0.08		0.04	0.07	0.13	0.11
<i>Sardinella hualiensis</i>	0.15	0.11	0.16		0.04	0.14	0.14
<i>Sardinella lemuru</i>	0.11	0.18	0.27	0.16		0.09	0.06
<i>Amblygaster sirm</i>	0.33	0.36	0.50	0.50	0.37		0.12
<i>Herklotsichthys quadrimaculatus</i>	0.30	0.37	0.39	0.47	0.25	0.47	
Control region							
Species	1	2	3	4	5		
<i>Sardinella gibbosa</i> Clade 1		0.07	0.08	0.07	0.06		
<i>Sardinella gibbosa</i> Clade 2	0.40		0.09	0.12	0.10		
<i>Sardinella fimbriata</i>	0.52	0.52		0.20	0.12		
<i>Sardinella hualiensis</i>	0.43	0.64	1.00		0.10		
<i>Sardinella lemuru</i>	0.40	0.61	0.75	0.66			
S7 intron							
Species	1	2	3	4	5	6	7
<i>Sardinella gibbosa</i> Clade 1		0.02	0.03	0.10	0.03	0.13	0.24
<i>Sardinella gibbosa</i> Clade 2	0.10		0.01	0.09	0.02	0.15	0.26
<i>Sardinella fimbriata</i>	0.10	0.03		0.11	0.02	0.17	0.27
<i>Sardinella hualiensis</i>	0.11	0.07	0.08		0.09	0.27	0.53
<i>Sardinella lemuru</i>	0.34	0.35	0.38	0.37		0.16	0.48
<i>Amblygaster sirm</i>	0.51	0.57	0.63	0.80	0.60		0.09
<i>Herklotsichthys quadrimaculatus</i>	0.62	0.72	0.77	1.00	0.95	0.41	

The standard error of mean is shown in the upper right diagonals.
doi:10.1371/journal.pone.0084719.t002

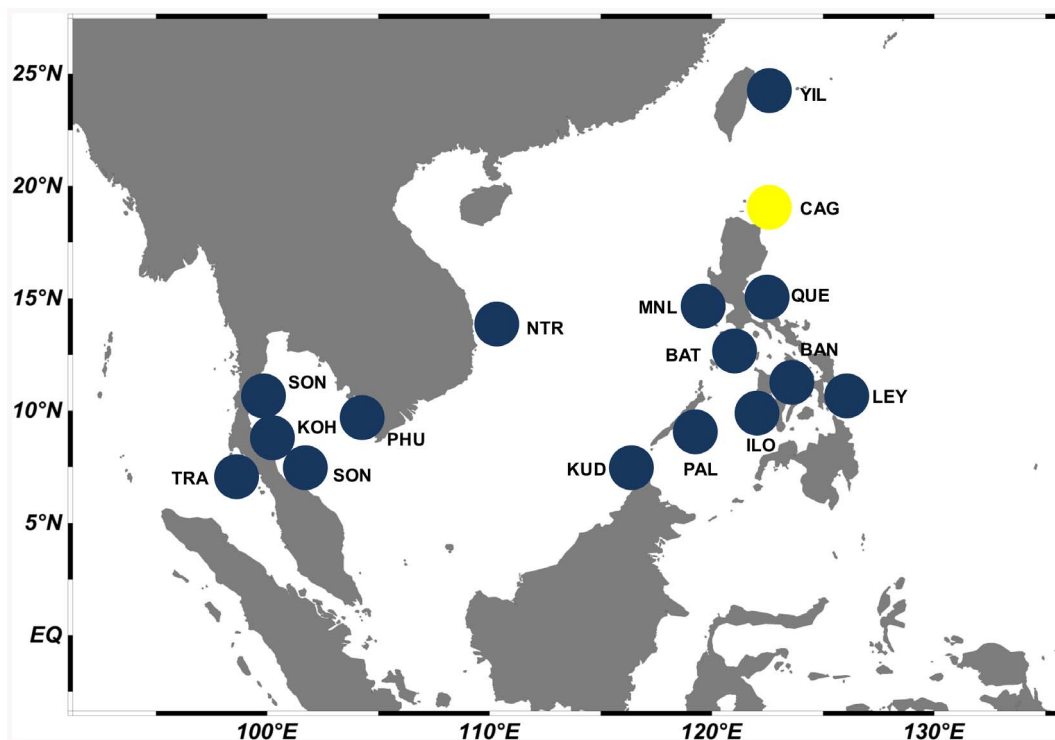


Figure 4. Sampling sites and clade distribution for *S. gibbosa*. Each color represents the cryptic clades. Clade 1 is represented in blue, clade 2 is shaded in yellow.

doi:10.1371/journal.pone.0084719.g004

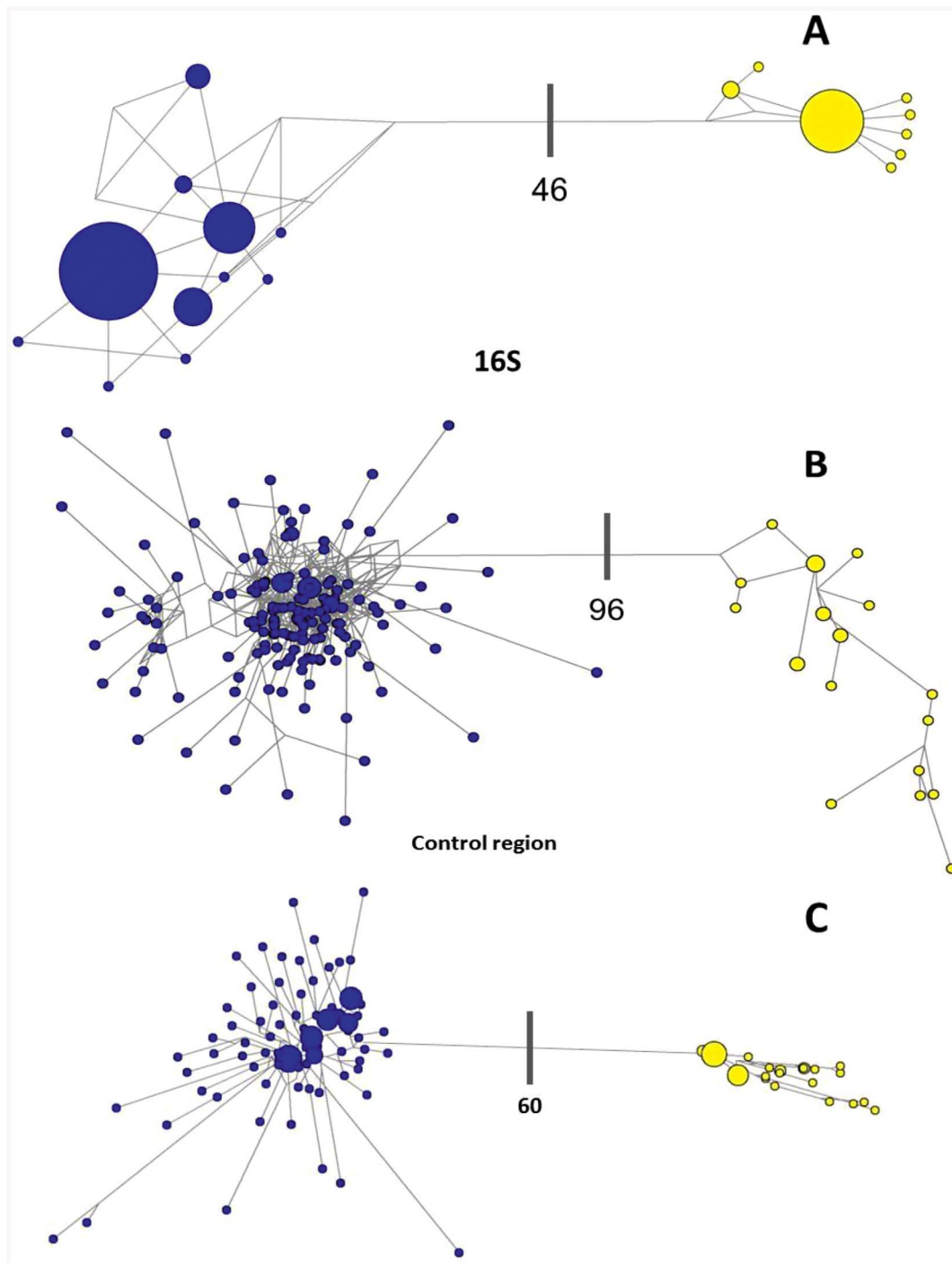


Figure 5. Median-joining network inferred for 16S (A) and control region (B) and S7. Each pie represents a haplotype proportional to its frequency. Thick bars represents step changes between Clade 1 (blue) and Clade 2 (yellow) which is labeled according to the number of base-pair difference.

doi:10.1371/journal.pone.0084719.g005

DNA polymerase. Thermal cycling conditions consisted of an initial denaturation of 94°C for 10 min followed by 38 cycles of DNA denaturing at 94°C for 30 s, primer annealing at 45°C for 45 s, and sequence extension at 72°C for 45 s, ending with a final extension of 72°C for 10 min. For the ribosomal S7 intron, we utilized PCR conditions as previously described [40]. Successful

PCR products were purified using ExoSAP-IT® (USB Corp, Cleveland, OH). The reaction mixture consisted of 2 µl of ExoSAP-IT and 22 µl of PCR product, and eventually incubated at 37°C for 15 min followed by another 15 min at 80°C to inactivate the enzyme. Purified PCR products were sent to either Macrogen, Inc. Korea or UC-Berkeley for DNA sequencing.

Sequence data was deposited on the public domain database GenBank [Accession numbers pending].

Phylogenetic reconstruction

Sequences were assembled in Geneious v5.4 [41] and aligned using MUSCLE v3.8.31 [42]. A best-fit nucleotide substitution model was determined using jMODELTEST v2 [43,44]. Phylogenetic analysis using maximum likelihood (ML) criteria was inferred from MEGA v5.2.1 [45] using the best-fit nucleotide substitution models, namely, Kimura-2-Parameter (K2P) for 16S, three parameter model (TPM) for control region and Hasegawa-Kishino-Yano (HKY) for the S7 intron. Also included in the analysis for outgroup comparison were the closely related taxa, namely, *Sardinella fimbriata*, *S. hualiensis*, *S. lemuru*, *Herklotsichthys quadrimaculatus* and *Amblygaster sim* sequences. Further, such species have overlapping geographic distribution with *S. gibbosa* throughout the Indo-West Pacific [23]. However, we excluded *Amblygaster sim* and *Herklotsichthys quadrimaculatus* as outgroups for control region dataset since they are highly divergent and inclusion of these taxa created large indels in sequence alignment. Allelic state of the nuclear S7 intron was estimated using PHASE v2.1 [46,47] as implemented in DnaSP v5.0 [48]. The phylogenetic network was inferred using the median-joining network implemented in NETWORK v4.6 [49] using the default settings.

Morphological analysis

To complement genetic data, variability within *S. gibbosa* from 10 individuals per site was quantified by morphometric measurements representing the head shape. Measurements (in mm) obtained using a Vernier caliper were standard length, snout length (tip of snout to eye), head length (tip of snout to edge of operculum), eye diameter (horizontal diameter), upper jaw length, and post-orbital length (right edge of eye to end of operculum). All measurements were converted into ratios to represent proportion with respect to standard length. A principal component analysis implemented in PC-ORD v4.10 [50] was performed on natural log-transformed ratios which separated morphological variations into linear combinations of variables that describe overall head shape. In addition, analysis of similarities (ANOSIM) and similarity of percentage analysis (SIMPER) were conducted on log-transformed morphometric ratios in Primer v5.2.4 [51] to determine the percentage contribution of morphometric ratios to the overall variations in head shape.

Results

Maximum-likelihood analysis of 16S rRNA sequences support the existence of two species within *S. gibbosa* (Figure 1). In concordance with 16S data, mitochondrial control region sequences revealed similar clustering (Figure 2). Clustering for both markers exhibited monophyletic clades with high bootstrap support. In addition, nuclear DNA sequences of the first intron of S7 gene revealed a deep divergence between Clade 1 and Clade 2 (Figure 3). None of the phylogenetic analyses indicated that the two morphotypes initially identified as *S. gibbosa* are sister species. Consistent across examined gene regions, genetic distances calculated for both mitochondrial and nuclear gene regions exhibited divergence comparable to species-level differentiation (Table 2). Clade 1 is broadly distributed across the collection sites except at the northernmost locations (Figure 4). In contrast, Clade 2 is geographically restricted to this one northernmost site

in Cagayan Province. Further, the single sample from Yilan County, Taiwan did not cluster with Clade 2, despite the site's close proximity with Cagayan Province. However, the current dataset for Taiwan, Vietnam, Thailand and Malaysia are only limited to the mitochondrial 16S gene region. Nevertheless, median joining network for all three markers, at least for Philippine sites, revealed numerous base-pair mutations between Clades 1 and 2 (Figure 5).

All specimens exhibited the diagnostic characters for *S. gibbosa*, including the dark spot at dorsal fin origin. However, head shape and pigmentation of both lower and upper jaws differ between the two clades identified using genetic markers (Figure 6). Principal component analysis (PCA) revealed strong morphometric differentiation in head dimensions (Figure 7). The first four principal components (PC) account for 95.42% of overall variance (PC1 – 48.74%; PC2 – 21.27%; PC3 – 14.94%; PC4 – 10.79%) (Table 3). PC1 was highly correlated with variance in upper jaw length, eye diameter, and post-orbital length, respectively. On the other hand, PC's 2 through 4 were associated with differences in the ratios of head length, upper jaw, and eye diameter. In concordance with the genetic clades, a surprisingly similar clustering was observed in plots of the principal components (Figure 7). Scatter plots for PC1 and PC3 separated collections from Quezon Province into a distinct cluster. Such grouping might indicate a unique sub-population or subspecies in Clade 1. A shorter snout and upper jaw with respect to head length in individuals from the four clades accounted for such clustering in principal component analysis.

Multivariate analysis of morphometric ratios using ANOSIM showed significant variations between head shape of the three

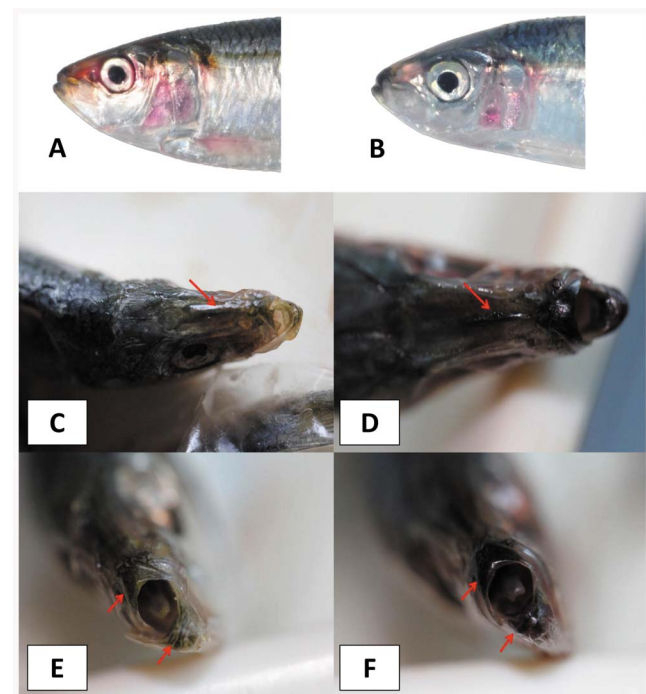


Figure 6. Photos of *Sardinella gibbosa* representing Clade 1 (A) and Clade 2 (B). Between-clade color differences in median frontal line of the head (C–D) and pigmentation of the upper and lower jaw (E–F).

doi:10.1371/journal.pone.0084719.g006

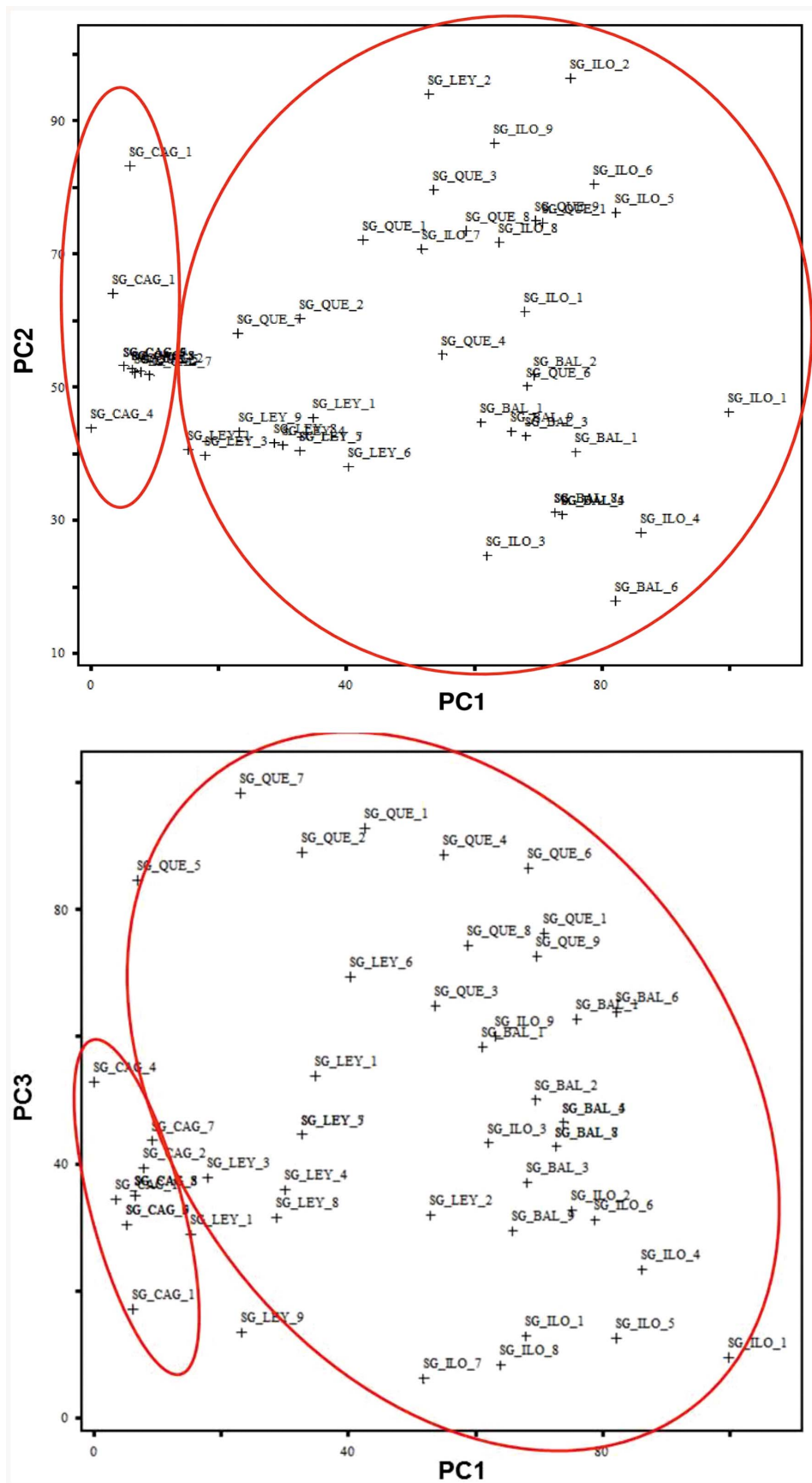


Figure 7. Plots of the first versus the second (A) and the first versus the third (B) principal components (PC) of 60 nominal specimens (n = 10 per population).
doi:10.1371/journal.pone.0084719.g007

genotypic clades ($R = 0.486$; $p < 0.01$). Differences in snout, post-orbital, and head lengths distinguished the two clades of *S. gibbosa* (Table 4). Strong differentiation in head and upper jaw length accounted for 52.26% of variation between Clade 1 and Clade 2. On the other hand, variations in eye, snout, and post-orbital length contributed 47.74% of overall difference between the two clades. Clade 1 had shorter heads with respect to standard length. Lastly, Clade 2 individuals had the lowest post-orbital length, and subsequently a shorter operculum.

In addition to morphometric difference, Clade 2 has a distinct black pigmentation on the edge of mouth and frontal line between the nostrils (Figure 6). Similar blackish coloration has been observed in the caudal fins of Clade 2. Further, a leaner body characterized Clade 2 in contrast with the more rounded shape of Clade 1 (data not shown). Lastly, Clade 2 lacked the gold stripe across the lateral body wall which characterized Clade 1 (Figure 6).

Discussion

Many discoveries of cryptic species have been based on prior observations of subtle behavioral, biological, or morphological intra-species variations [5]. However, phenotypic differentiation does not necessarily complement genotypic divergence [9], as evident in the lack of congruence between genetics and diagnostic morphological characters [6,10,52]. In extreme cases, morphological variations are randomly shared among genetically distinct lineages within a cryptic species complex [11,53]. To avoid the inconsistency between genetics and morphology, the straightforward approach is to identify cryptic species using multi-locus genetic data [54,55], a method that can help avoid the pitfalls of morphological species delimitation [12]. Cryptic species identified through genetic clustering can then be bolstered by support from additional morphological or biological traits.

In *S. gibbosa*, molecular evidence from both mitochondrial and nuclear DNA strongly supports two cryptic species. Clade 1 is widely distributed throughout the Philippine islands, while Clade 2 is geographically restricted to the Cagayan Province. Such allopatric distribution has been reported to occur in other cryptic species [15,16]. In addition, the two clades exhibited the same clustering for all markers (Figures 1–3), a finding consistent with

the ‘genotypic-clustering’ species definition [12], and substantiated by agreement between multi-locus genotypic data [56,57]. Likewise, a lack of reciprocal monophyly in Clades 1 and 2 showed that the two lineages are different and not sister species. Clades 1 and 2 were paraphyletic with each other, a phylogenetic pattern that commonly occurs among cryptic species [58,59]. In addition, the 10–40% genetic distances between cryptic clades are comparable to species-level differences (Table 2) [60]. In some pairs, genetic distances for 16S rRNA and S7 intron of *S. gibbosa* exceed levels typically distinguishing closely-related species [60]. It is also interesting to note that Clades 1 and 2 do not have shared haplotypes in both the conserved 16S rRNA and the polymorphic mitochondrial control region (Figure 5). Consistent patterns in both maternally and bi-parentally inherited genetic markers demonstrate a lack of gene flow between the two cryptic species. Such patterns fall within the framework of the general species concept for two unique species [1]. Hence, genetic information from both mitochondrial and nuclear DNA presents solid evidence for two biologically distinct species.

Distinct morphometric variations in head shape characterized both Clade 1 and 2 of the *S. gibbosa* species. Multivariate analysis of head measurements revealed clustering comparable to the genetic clades. Similar clustering due to head shape variations have characterized sub-species within a sardine [23,61], a result later confirmed by molecular evidence from mitochondrial data [62]. Morphological differences between closely related sardines are often characterized by slight differences in measurements or meristic counts, resulting in an ambiguous and often controversial taxonomic status [23]. For instance, the sister sardine species *Sardinella tawilis* and *Sardinella huailiensis* share diagnostic characters; and excluding habitat preference, differ only in head length and lower gillraker count [63]. However, intra-species morphological variations in sardines are often presumed to be an artifact of localized adaptations to environment, due to a lack of support from significant genetic differentiation between morphological forms [64,65]. In contrast, the morphological disparity between the cryptic clades of *S. gibbosa* complements genetic divergence (Figure 6), and thus is not a mere localized ecological adaptation. Lastly, clustering in PCA due to head shape in Clades 1 and 2 of *S. gibbosa* falls within the phenetic-clustering species delimitation, as

Table 3. Summary of principal component analysis (eigenvalues and eigenvectors) calculated from 5 morphometric ratios describing head shape.

	PC 1	PC 2	PC 3	PC 4	PC 5
Eigenvalue	2.440	1.063	0.747	0.540	0.213
% variance	48.74	21.27	14.94	10.79	4.27
Eigenvector					
head	0.403	−0.126	0.875	0.101	−0.213
snout	−0.130	0.893	0.228	−0.366	−0.011
eye	−0.515	0.123	0.291	0.651	0.459
postorbital	−0.453	−0.408	0.306	−0.652	0.331
upper jaw	−0.591	−0.077	0.058	0.082	−0.796

Loadings with absolute values > 0.3 are shown in bold.

doi:10.1371/journal.pone.0084719.t003

Table 4. Summary of 5 morphometric measurements representing variations in head shape.

Morphometric measurement	Morphometric ratio (% Ls)		Clade 1 vs. Clade 2 (% difference)	
	Clade 1	Clade 2	Contribution	Cumulative
head length	22.33	24.27	32.00	32.00
upper jaw	40.48	34.62	20.25	52.26
eye	26.82	23.08	18.78	71.04
snout	30.31	30.06	15.74	86.78
post-orbital	40.58	38.46	13.22	100.00

Morphometric ratio refers to each character value as a percentage of standard length (% Ls). Contribution and cumulative differences calculated using SIMPER describes, in percentage, each morphometric character's contribution to head shape variation between Clade 1 and Clade 2.

doi:10.1371/journal.pone.0084719.t004

there is a lack of intermediate forms in between the two clades [12,66].

Combined, genetic and morphological data reveal a hidden diversity in a common and commercially important sardine. Our findings expand the previous investigations on the biology, ecology, and morphology of *S. gibbosa* that alluded to a cryptic diversity [25,31]. Discovery of new fish species in the Northern Philippines [67,68] including a sardine beyond its previously known distribution [69], suggests that this region harbors undocumented and unique fauna. Such a pattern presents the possibility that Clade 2 might be a new species. Alternatively, Clades 1 and 2 might be previously documented synonyms of *S. gibbosa* [23]. Based on geographic proximity and morphological similarity, the most likely candidate synonym is *S. taiwanensis* [35]; however, further scrutiny of type specimens is necessary for validation. Nevertheless, the findings in this study demonstrate that a combination of both morphological and genetic data is essential to assess diversity in taxonomically ambiguous sardines. Here, strong evidence of two ecologically similar, but genetically and morphologically distinct species warrants

appropriate management strategies for separate sardine fisheries.

Acknowledgments

The authors thank the researchers at Genetic Fingerprinting Laboratory of NFRDI for their superb support. Sample collection was done with the assistance of the regional offices of Philippine Bureau of Fisheries and Aquatic Resources through the National Stock Assessment Program. Samples were collected and verified with the help of Eunice Bognot of BFAR-NFRDI. Additional feedbacks on the paper were contributed by Dr. Mary Beth Maningas of the University of Santo Tomas. We would like to extend our utmost thanks to Benedict Maralit and Sheryll Hipolito of Tokyo University of Marine Science and Technology, for gathering archived literature from Japan. Also, we would like to recognize Mr. Fumiaki Nishihara for his excellent work on translating Japanese texts.

Author Contributions

Conceived and designed the experiments: RCT MDS DAW. Performed the experiments: RCT. Analyzed the data: RCT DAW. Contributed reagents/materials/analysis tools: MDS DAW RCT KEC. Wrote the paper: RCT DAW MDS KEC.

References

- Haudorf B (2011) Progress towards a general species concept. *Evolution* 65: 923–931.
- De Quieroz K (2007) Species concepts and species delimitation. *Systematic Biology* 56: 879–886.
- Hey J, Waples R, Arnold M, Butlin R, Harrison R (2003) Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology and Evolution* 18:597–603.
- Mayr E (1963) Animal species and evolution. Oxford University Press, London.
- Bickford D, Lohman D, Sodhi N, Ng P, Meier R, et al. (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* 22: 148–155.
- Knowlton N (1993) Sibling species in the sea. *Annual Review of Ecology and Systematics* 24: 189–216.
- Piggott M, Chao N, Beheregaray L (2011) Three fishes in one: cryptic species in an Amazonian floodplain forest specialist. *Biological Journal of the Linnean Society* 102: 391–403.
- Boissin E, Feral J, Chenuil A (2008) Defining reproductively isolated units in a cryptic and syntopic species complex using mitochondrial and nuclear markers: the brooding brittle star, *Amphipholis squamata* (Ophiuroidea). *Molecular Ecology* 17: 1732–1744.
- Knowlton N (2000) Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* 420: 73–90.
- Carolan J, Murray T, Fitzpatrick U, Crossley J, Schmidt H, et al. (2012) Colour patterns do not diagnose species: quantitative evaluation of a DNA barcoded cryptic bumblebee complex. *PLOS-ONE* 7: e29251.
- Hebert P, Penton E, Burns J, Janzen D, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgator*. *Proceedings of the National Academy of Sciences USA* 101: 14812–14817.
- Mallet J (1995) A species definition for the modern synthesis. *Trends in Ecology and Evolution* 10: 294–299.
- Beheregaray L, Caccone A (2007) Cryptic biodiversity in a changing world. *Journal of Biology* 6: 9.
- Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographic regions. *BMC Evolutionary Biology* 7: 121.
- Ross K, Gotzek D, Ascunce M, Shomeaker D (2010) Species delimitation: a case study in a problematic ant taxon. *Systematic Biology* 59: 162–184.
- Brown D, Brenman R, Koepfli K, Pollinger JP, Mila B, et al. (2007) Extensive population genetic structure in the giraffe. *BMC Biology* 5: 57.
- Stuart B, Inger R, Voris H (2006) High level of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. *Biology Letters* 2: 470.
- Goodman S, Maminirina C, Weyeneth N, Bradman H, Christidis L, et al. (2009) The use of molecular and morphological characters to resolve the taxonomic identity of cryptic species: the case of the *Miniopterus manavi* (Chiroptera, Miniopteridae). *Zoologica Scripta* 38: 339–363.
- Hubert N, Meyer C, Bruggemann H, Guerin F, Komono R, et al. (2012) Cryptic diversity in Indo-Pacific coral-reef fishes revealed by DNA-barcoding provides new support to the centre-of-overlap hypothesis. *PLOS ONE* 7: e28987.
- Trontelj P, Fiser C (2009) Cryptic species diversity should not be trivialized. *Systematics and Biodiversity* 7:1–3.
- Atarhouch T, Ruber L, Gonzales E, Albert E, Rami M, et al. (2006) Signature of an early genetic bottleneck in a population of Moroccan sardines (*Sardina pilchardus*). *Molecular Phylogenetics and Evolution* 39: 373–383.
- FAO. © 2006–2013. National Aquaculture Sector Overview. Philippines. National Aquaculture Sector Overview Fact Sheets. Text by Paclibare, J.O. In: *FAO Fisheries and Aquaculture Department* [online]. Rome. Updated 1 February 2005. [Cited 13 February 2013].
- Whitehead P (1985) Clupeoid fishes of the world (Suborder Clupeoidei): An annotated & illustrated catalogue of the herrings, sardines, pilchards, sprats, shads, anchovies & wolf-herrings. *FAO Fisheries Synopsis* vol 7, Rome.

24. Lavoue S, Miya M, Musikasinthorn P, Chen W, Nishida M (2012) Mitogenomic evidence for an Indo-West Pacific origin of the Clupeoidei (Teleostei: Clupeiformes). *PLoS ONE* 8(2): e56485.
25. Okera W (1974) Morphometrics, 'condition' & gonad development of the East African *Sardinella gibbosa* (Bleeker) & *Sardinella albella* (Valenciennes). *Journal of Fish Biology* 6: 801–812.
26. Nair R (1973) Indian sardines (their biology & fishery). CSIR Zoological Monograph (2). CSIR, New Delhi.
27. Willette D, Bognot E, Mutia M, Santos M (2011) Biology and ecology of sardines in the Philippines: a review. *BFAR Technical Paper Series* 13: 1–20.
28. Nyunja J, Mavuti K, Wakwabi E (2002) Trophic ecology of *Sardinella gibbosa* (Pisces: Clupeidae) & *Atherinomorus lacunosus* (Pisces: Atherenidae) in Mtwapa Creek & Wasini Channel, Kenya. *Western Indian Ocean Journal of Marine Science* 1: 181–189.
29. Lazarus S (1975) Observations on the food & feeding habits of *Sardinella gibbosa* from Vizhinjam. *Indian Journal of Fisheries* 24: 107–112.
30. Sekharan K (1969) Spawning concentrations of the sardine, *Sardinella gibbosa* (Bleeker), off the north Andhra coast in March–April 1969. *Indian Journal of Fisheries* 16: 156–160.
31. Dutt S (1961) Biometric studies on *Sardinella* spp. of Waltair coast. *J Zoo Soc India* 13: 78–89.
32. Okera W (1970) An Analysis of the features of *Sardinella gibbosa* (Bleeker) scales, with special reference to the problem of age determination. *Journal of the East Africa Natural History Society and National Museum* 28: 1–14.
33. Kishinouye K (1908) Notes on the natural history of the sardine (*Clupea melanosticta* Schlegel). *Journal of the Imperial Fisheries Bureau* 14(3): 71–105, Pls. 13–21.
34. Whitley G (1940) Illustrations of some Australian fishes. *Australian Zoologist* 9: 397–428m Pls. 30–31.
35. Raja B, Hiyama Y (1969) Studies on the systematics and biometrics of a few Indo-Pacific sardines. *Records of Oceanographic Works in Japan* 10(1): 75–103.
36. Walsh P, Metzger A, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR based typing from forensic material. *Biotechniques* 10: 506–513.
37. Bruford M, Hanotte O, Brookfield J, Burke T (1998) Multilocus and single-locus DNA fingerprinting. In: *Molecular Genetic Analysis of Populations: A Practical Approach* 2nd ed, (ed Hoelzel A), pp. 287–336. IRL Press, Oxford, UK.
38. Palumbi S (1996) Nucleic acids II: the polymerase chain reaction. In *Hills DM, Moritz C, Mable BK (eds) Molecular Systematics*, Sinauer, Sunderland.
39. Lee W, Howell W, Kocher T (1995) Structure & evolution of teleost mitochondrial control regions. *Journal of Molecular Evolution* 41: 54–66.
40. Chow S, Hazama K (1998) Universal PCR primers for S7 ribosomal protein gene introns in fish. *Molecular Ecology* 7: 1247–1263.
41. Geneious version 5.4 created by Biomatters. Available: <http://www.geneious.com>
42. Edgar R (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acid Research* 32 (5): 1792–1797.
43. Darriba D, Taboada G, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9: 772.
44. Guindon S, Gascuel O (2003) A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* 52: 696–704.
45. Tamura K, Peterson D, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum-likelihood, evolutionary distance and maximum parsimony method. *Molecular Biology and Evolution* 28: 2731–2739.
46. Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics* 73: 1162–1169.
47. Stephens M, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 68: 978–989.
48. Librado P, Rozas J (2009) DnaSP v5: As software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
49. Bandelt H, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37–48.
50. McCune B, Mefford M (1999) Multivariate analysis of ecological data v4.10. MjM Software, Gleneden Beach, Oregon.
51. Clarke K, Warwick R (2001) Change in marine communities: an approach to statistical analysis and interpretation. Primer-E Ltd. Plymouth.
52. Carolan J, Murray T, Fitzpatrick U, Crossley J, Schmidt H, et al. (2012) Colour patterns do not diagnose species: Quantitative evaluation of a DNA barcoded cryptic bumblebee complex. *PLoS-ONE* 7(1): e29251.
53. Colborn J, Crabtree R, Shaklee J, Pfeiler E, Bowen B (2001) The evolutionary enigma of bonefishes (*Albula spp.*): cryptic species and ancient separations in a globally distributed shorefish. *Evolution* 55: 807–820.
54. Matthews L, Anker A (2009) Molecular phylogeny reveals extensive ancient and ongoing radiations in a snapping shrimp species complex (Crustacea, Alpheidae, *Alpheus armillatus*). *Molecular Phylogenetics and Evolution* 50: 268–281.
55. Ladner J, Palumbi S (2012) Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes. *Molecular Ecology* 21: 2224–2238.
56. Niemiller M, Near T, Fitzpatrick B (2012) Delimiting species using multilocus data: diagnosing cryptic diversity in the southern cavefish, *Typhlichthys subterraneus* (Teleostei: Amblyopsidae). *Evolution* 66–3: 846–866.
57. Dupuis J, Roe A, Sperling F (2012) Multilocus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology* 21: 4422–4436.
58. Shaffer H, Thomson R (2007) Delimiting species in recent radiations. *Systematic Biology* 56: 896–907.
59. Funk D, Omland K (2003) Species-level paraphyly and polyphyly: Frequency, causes and consequences, with insights from animal mitochondrial DNA. *Annual Reviews in Ecology and Systematics* 34: 397–423.
60. Lavoue S, Sullivan J, Hopkins C (2003) Phylogenetic utility of the first two introns of the S7 ribosomal protein gene in African electric fishes (Mormyroidea: Teleostei) and congruence with other molecular markers. *Biological Journal of the Linnean Society* 78: 273.
61. Silva A (2003) Morphometric variation among sardine (*Sardina pilchardus*) populations from the northeastern Atlantic and the western Mediterranean. *ICES J Mar Sci* 60: 1352–1360.
62. Atarhouch T, Ruber L, Gonzales E, Albert E, Rami M, et al. (2006) Signature of an early genetic bottleneck in a population of Moroccan sardines (*Sardina pilchardus*). *Molecular Phylogenetics and Evolution* 39: 373–383.
63. Willette D, Carpenter K, Santos M (2014) Evolution of the freshwater sardinella, *Sardinella tawilis* (Clupeiformes: Clupeidae), in Lake Taal, Philippines and identification of its marine sister-species, *Sardinella huailiensis*. *Bulletin of Marine Science* 90: In press.
64. Kinsey S, Orsoy T, Bert T, Mahmoudi B (1994) Population structure of the Spanish sardine *Sardinella aurita*: natural morphological variation in a genetically homogenous population. *Marine Biology* 118: 309–317.
65. Hedgecock D, Hutchinson E, Li G, Sly F, Nelson K (1989) Genetic and morphometric variation in the Pacific sardine, *Sardinops sagax caerulea*: Comparisons and contrasts with historical data and with variability in the Northern anchovy, *Engraulis mordax*. *Fishery Bulletin* 87: 653–671.
66. Sokal R, Crovello T (1970) The biological species concept: A critical evaluation. *The American Naturalist* 104: 127–153.
67. Gill A, Williams J (2011) Description of two new species of Pseudochrominae from northern Palawan and Mindoro, Philippine Islands (Teleostei: Perciformes: Pseudochromidae). *Zootaxa* 3140: 49–59.
68. Gill A, Shao-K, Chen J (1995) *Pseudochromis cerasina*, a new species of pseudochromine dotyback fish from the West Pacific (Perciformes: Pseudochromidae). *Proceedings of the Biological Society of Washington* 117:17–22.
69. Willette D, Santos M, Aragon M (2011) First report of the Taiwan sardinella *Sardinella huailiensis* (Clupeiformes: Clupeidae) in the Philippines. *Journal of Fish Biology* 79: 2087–2094.